



FT Modified-site /note= "potential protein kinase C phosphorylation site"  
 FT 121  
 FT /note= "potential casein kinase II phosphorylation site"  
 FT 130  
 FT Modified-site /note= "potential casein kinase II phosphorylation site"  
 FT 162..218  
 FT Domain /note= "LIM domain"  
 FT 175  
 FT Modified-site /note= "potential casein kinase II or protien  
 FT 201 kinase C phosphorylation site"  
 FT Modified-site /note= "potential casein kinase II or protien  
 FT 202 kinase C phosphorylation site"  
 FT Modified-site /note= "potential casein kinase II phosphorylation site"  
 FT 221..275  
 FT Domain /note= "LIM domain"  
 FT 223  
 FT Modified-site /note= "potential protein kinase C phosphorylation site"  
 FT 238  
 FT Modified-site /note= "potential casein kinase II phosphorylation site"  
 FT 238  
 XX WO200014231-A1.  
 PN  
 XX 16-MAR-2000.  
 XX  
 XX 01-SEP-1999; 99WO-US20239.  
 XX  
 XX 03-SEP-1998; 98US-0155260.  
 PR  
 XX  
 XX (INCY-) INCYTE PHARM INC.  
 PA  
 XX Tang YT, Guegler KJ, Corley NC, Patterson C;  
 PI  
 XX WPI; 2000-256982/22.  
 XX  
 XX N-PSDB; AAZ99798.  
 DR  
 PT A purified polypeptide designated LDPH (LIM domain protein homolog),  
 PT useful for treating, preventing, and diagnosing various cancers,  
 PT reproductive disorders and autoimmune disorders.  
 XX  
 PS Claim 1; Page 52-53; 59pp; English.  
 XX  
 CC The present sequence represents a human LIM domain protein homologue.  
 CC The LIM domain is so named because it was first described in three  
 CC proteins from Drosophila melanogaster, designated L, I, and M. The LIM  
 CC domain is rich in cysteine. Nucleic acids encoding the LIM protein  
 CC homologue were first identified in incyte clone 1925017 from a breast  
 CC tissue cDNA library. The LIM polynucleotide and polypeptide are  
 CC useful for treating, preventing, and diagnosing various cancers,  
 CC reproductive disorders and autoimmune disorders.  
 XX  
 SQ Sequence 280 AA;  
 Query Match 100.0%; Score 1635; DB 21; Length 280;  
 Best Local Similarity 100.0%; Pred.No. 1.4e-127;  
 Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSESFDCAKNEISLYGRKVIQTDSPGVCPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60  
 DB 1 MSESFDCAKNEISLYGRKVIQTDSPGVCPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60  
 QY 61 FHEGCFRCRCQRLADEPFTQDSPELLCNDVCYSAFSSQCSAGETVMPGSRKLEYGGQ 120  
 DB 61 FHEGCFRCRCQRLADEPFTQDSPELLCNDVCYSAFSSQCSAGETVMPGSRKLEYGGQ 120  
 QY 121 TWHEHCFLCSCGEPLGSRSPVDPKGAHYCVPCYENKFAPCARCSKTLTQGGVYTRDQP 180  
 DB 121 TWHEHCFLCSCGEPLGSRSPVDPKGAHYCVPCYENKFAPCARCSKTLTQGGVYTRDQP 180  
 QY 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSSCKRPIVGLGGKYYVSFE 240  
 DB 181 WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGELFAPKCSSCKRPIVGLGGKYYVSFE 240

QY 241 DRHHHNCFCARGSTSLVGQGFVDPDGVLCQCSQAGP 280  
 DB 241 DRHHHNCFCARGSTSLVGQGFVDPDGVLCQCSQAGP 280  
 RESULT 2  
 ABP64881  
 ID ABP64881 standard; Protein; 284 AA.  
 XX  
 AC ABP64881;  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human protein SEQ ID 541.  
 XX  
 KW Human; expressed sequence tag; EST;  
 KW haematopoietic disorder; central nervous system disease; viral infection;  
 KW peripheral nervous system disease; non-healing wound; infectious disease;  
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;  
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;  
 KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
 KW immunostimulant; cerebroprotective.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200259260-A2.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 16-NOV-2001; 2001WO-US42950.  
 XX  
 PR 17-NOV-2000; 2000US-0714936.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX  
 XX WPI; 2002-590824/63.  
 DR N-PSDB; ABQ99467.  
 DR  
 XX New isolated polynucleotide, useful in research, diagnostic or  
 PT therapeutic methods, e.g. preventing or treating disorders involving  
 PT aberrant protein expression or biological activity.  
 XX  
 PS Claim 20; SEQ ID 541; 394pp; English.  
 CC  
 CC The present invention relates to novel human coding sequences  
 CC (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are  
 CC useful in therapeutic, diagnostic and research methods. The  
 CC polynucleotides may be used in the field of molecular biology as  
 CC hybridisation probes, primers for PCR, for chromosome and gene mapping,  
 CC for the recombinant production of protein, or in generation of anti-sense  
 CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed  
 CC sequence tags (ESTs) for identifying expressed genes or for physical  
 CC mapping of the human genome. The proteins may be used as molecular weight  
 CC markers, or as nutritional sources or supplements. The proteins may be  
 CC used to maintain and expand cell population in a totipotential or  
 CC pluripotential state useful for re-engineering damaged or diseased  
 CC tissues, transplantation, manufacture of bio-pharmaceuticals or the  
 CC development of bio-sensors. The polynucleotides and proteins are useful  
 CC for preventing, treating or ameliorating disorders involving aberrant  
 CC protein expression or biological activity, e.g. haematopoietic disorders,  
 CC central/peripheral nervous system diseases, mechanical and traumatic  
 CC disorders, non-healing wounds, immune deficiencies and disorders,  
 CC infectious diseases caused by viral, bacterial or fungal infection,  
 CC autoimmune disorders, allergic reactions and conditions, coagulation  
 CC disorders, or cancer. The polynucleotide sequences of the invention were  
 CC assembled from ESTs isolated mainly by sequencing by hybridisation, and  
 CC in some cases, sequences obtained from one or more public databases.  
 CC Note: The sequence data for this patent did not form part of the printed







PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231442.  
PR 08-SEP-2000; 2000US-0231443.  
PR 08-SEP-2000; 2000US-0231444.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
DR WPI; 2001-465570/50.  
DR N-PSDB; AAL01353.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen -  
XX is used in preventing, treating or ameliorating a medical condition -  
XX  
XX Claim 11; SEQ ID NO 4041; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
XX number of human reproductive system related antigens, these can be used  
XX in the prevention and treatment of reproductive system disorders,  
XX including cancer. The present sequence is a protein of the invention.  
XX  
XX Sequence 245 AA;  
Query Match 43.8%; Score 716; DB 22; Length 245;  
Best Local Similarity 46.2%; Pred. No. 1.5e-51;  
Matches 109; Conservative 48; Mismatches 79; Indels 0; Gaps 0;  
QY 40 CAECQOLIGHDSRELFVEDRHFHGGCFRCRCORSIADEPFTQDSSELLCNDYCYSFSS 99  
Db 2 CEECKPIESDSKDLCKDRHWHGECFKCTKNHSLVEKPFPAKDERLLCTECSNECSS 61  
QY 100 QCSACGETVMPGSRKLEYGGQTWHEHCFLSCGCEQPLGSRSFVDPKGAHYCVPCYENKFA 159  
Db 62 KCFCHKRTIMPGSRKMEFGNYWHXTCXVCENCRCQPIXTKPLISKESGNYCVPCXEKEFA 121  
QY 160 PRCARCSKLTGGVTYRQPHRECLVCTGCTCOTPLAGQOFTSRDEDPYCVACFGLFAP 219  
Db 122 HYCNFKKVTJSGGITFCQDLQWHKEXFLCSGCRDRDCEQFMGRDDYPPFCXDCYNHYLAN 181  
QY 220 KCSSCKRPVGLGGKGYVSPEDRHHNCFSCARCSTSLVGQGFVPDQVLCQGC 275  
Db 182 KCVACSKPISGLTGANKFICFQDSQWHSECFNCCKCSVSLVGKGLTQNKEIFCQKC 237  
RESULT 5  
AAU20437  
ID AAU20437 standard; Protein; 245 AA.  
XX  
XX AAU20437;  
XX  
XX DT 06-DEC-2001 (first entry)  
XX





CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 328 AA;  
Query Match 25.3%; Score 413.5; DB 22; Length 328;  
Best Local Similarity 27.5%; Pred. No. 2.3e-26;  
Matches 84; Conservative 48; Mismatches 135; Indels 39; Gaps 6;  
QY 3 ESFDCAKNESLYGKRYQTDSGPCYPCYDNTFANTCAEQQLIGHDSRELFDHGFH 62  
Db 29 EHFLCHHDEQILDATFNVSQGEPCVNCFCFERYTYTCAGCKKPI--LEKTCAMGESWH 86  
QY 63 EGCFFRC--RCQSLADEPFTCCQDSELLCNDYCSAFSSQCSAGCTVMPGS----- 112  
Db 87 EDCP-CCGACKKPLANFTFYERDCKPYCKDYEDLFAARCAKCEKPTDTSAVLAMNVKW 145  
QY 113 -----RKLEYGGQTWHEHCFLCSCGEQPLGSRSFYDCKGAHYCV 151  
Db 146 HRDCFRNCNIVCHKQEAITKRMITALGKTWHEPHEFLCHHDEQILDATFNVSQGEPCV 205  
QY 152 PCYENKFAPRCAKSKTLTQGVYRQDPWHRECLVCTG-CQPLAGQOFTSRDEDPYCV 210  
Db 206 KCFERYTYTCAGCKKPILEKTCAMGERWHEACFCGAGACKKPLASQTFYERDCKPYCK 265  
QY 211 ACFGELFAPKSSCKRPITVGLGGKYVSFEDRHHNCFSCARCTSLVGGQFVDPGQV 270  
Db 266 QDYENLFAARCAKCEKPI-----DSAVLAMNVKHNRCFCQNCENPITSQTFIDGKP 321  
QY 271 LQCGCS 276  
Db 322 VCPACN 327  
RESULT 9  
ABG11960  
ID ABG11960 standard; Protein; 116 AA.  
XX  
AC ABG11960;  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #11951.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR N-PSDB; AAS76147.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20; SEQ ID NO 42319; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 116 AA;  
Query Match 21.3%; Score 349; DB 22; Length 116;  
Best Local Similarity 48.6%; Pred. No. 1.7e-21;  
Matches 52; Conservative 23; Mismatches 32; Indels 0; Gaps 0;  
QY 169 LTQGVYRQDPWHRECLVCTGCOTPLAGQOFTSRDEDPYCVACGELFAPKSSCKRP 228  
Db 2 ITSGGITFCQDLWHKECFCLSGCRKDLCEQFMGRDDYPCVVCYNHLYANKVACSCKPI 61  
QY 229 VGLGGKYVSPEDRHHNCFSCARCTSLVGGQFVDPGDOVLQCGC 275  
Db 62 SGLTGAKFICFQDSQWHSECFNCCKCSVSLVGKGLTQNKEIFCQKC 108  
RESULT 10  
ABB67068  
ID ABB67068 standard; Protein; 746 AA.  
XX  
AC ABB67068;  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 27996.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL11171.  
XX

















```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0460 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 314 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1516012
;
US-09-008-465-3
Query Match 18.7%; Score 306.5; DB 3; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDYDRHFH 62
DQ 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGEFT--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117
DQ 92 PECFRCDLCQEVLDIGFVKNAGRHLRCRPNHREKARGLGKGYICOKCHAIL--DQPLIF 149
QY 118 GGQTWHEHCFILCSGCEQPL--GSRSFVPDKGAHYCVPCYENKFAPRCARCSTLTQGGVT 175
DQ 150 KNDPYHPDHFNCANGKELTADAREL---KGELYCLPCHDKMGVPICGACRRPIEGRVYN 206
QY 176 YRDQWHRRECLVCTGCTPLAGQQFTSRDEDPYCVACFGEFAPKCSSCKRPITVGLGGK 235
DQ 207 AMGKQWHEHFVCAKCEKPFGLGHRHYERKGLAYCETHYNQLFGDVCFFHCNRYI----EGD 262
QY 236 YVSFEDRHHNCFSCARCSTSL-VGQGFVPDGDQVLCQGC 275
DQ 263 VVSALNKAWCNCFACSTCNTKLTILKNKFEVDFMDKPVCKKC 303

RESULT 5
US-09-276-851-2
; Sequence 2, Application US/09276851
; Patent No. 6245522
; GENERAL INFORMATION:
; APPLICANT: Rearden, Ann
; TITLE OF INVENTION: A MASTER MOLECULAR RHEOSTAT SWITCH FOR
; TITLE OF INVENTION: CELL SIGNALING
; FILE REFERENCE: REGEN1440
; CURRENT APPLICATION NUMBER: US/09/276,851
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-276-851-2
Query Match 18.7%; Score 306.5; DB 3; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDYDRHFH 62
DQ 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGEFT--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117
DQ 92 PECFRCDLCQEVLDIGFVKNAGRHLRCRPNHREKARGLGKGYICOKCHAIL--DQPLIF 149
QY 118 GGQTWHEHCFILCSGCEQPL--GSRSFVPDKGAHYCVPCYENKFAPRCARCSTLTQGGVT 175
DQ 150 KNDPYHPDHFNCANGKELTADAREL---KGELYCLPCHDKMGVPICGACRRPIEGRVYN 206
QY 176 YRDQWHRRECLVCTGCTPLAGQQFTSRDEDPYCVACFGEFAPKCSSCKRPITVGLGGK 235
DQ 207 AMGKQWHEHFVCAKCEKPFGLGHRHYERKGLAYCETHYNQLFGDVCFFHCNRYI----EGD 262
QY 236 YVSFEDRHHNCFSCARCSTSL-VGQGFVPDGDQVLCQGC 275
DQ 263 VVSALNKAWCNCFACSTCNTKLTILKNKFEVDFMDKPVCKKC 303

RESULT 6
US-09-528-959-3
; Sequence 3, Application US/09528959
; Patent No. 6379904
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Corley, Neil C.
; TITLE OF INVENTION: HUMAN PINCH PROTEIN HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/528,959
FILING DATE: 20-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/008,465
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0460 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1516012
;
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-528-959-3
Query Match 18.7%; Score 306.5; DB 4; Length 314;
Best Local Similarity 27.4%; Pred. No. 2.8e-21;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

QY 3 ESFDCAKNESLYGRKYIQTDSGPYCPYNTFANTCAEQQLIGHDSRELFDYDRHFH 62
DQ 34 QCFVCAQCFOQFPPEGLFVEFEGRKYCEHDFQMLFAPCCHQCGEFT--IGRVIKAMNNSWH 91
QY 63 EGCPRCCRCQSLADEPTCQDSSELLCNDYCYSAFSSQ-----CSACGETVMPGSRKLEY 117
DQ 92 PECFRCDLCQEVLDIGFVKNAGRHLRCRPNHREKARGLGKGYICOKCHAIL--DQPLIF 149
QY 118 GGQTWHEHCFILCSGCEQPL--GSRSFVPDKGAHYCVPCYENKFAPRCARCSTLTQGGVT 175
DQ 150 KNDPYHPDHFNCANGKELTADAREL---KGELYCLPCHDKMGVPICGACRRPIEGRVYN 206
QY 176 YRDQWHRRECLVCTGCTPLAGQQFTSRDEDPYCVACFGEFAPKCSSCKRPITVGLGGK 235
DQ 207 AMGKQWHEHFVCAKCEKPFGLGHRHYERKGLAYCETHYNQLFGDVCFFHCNRYI----EGD 262
QY 236 YVSFEDRHHNCFSCARCSTSL-VGQGFVPDGDQVLCQGC 275
DQ 263 VVSALNKAWCNCFACSTCNTKLTILKNKFEVDFMDKPVCKKC 303

```

Db 92 PECFRCDLQCEVLADIGFVKNAGRHLCRPNCHNREKARGLGKVICQKCHAI--DEOPLIF 149  
QY 118 GGTWHEHCFACGCGQPL--GSRFVDPKGAHYVPCYENKFAPCARCCKTTLTGGVYT 175  
Db 150 KNDPYHDFHCANCCKELTADAREL---KGELYCLPCHDKWGPICGACRPIEGRVYN 206  
QY 176 YRDQPHRECLVCTGCTPLAGQOFTSRDEDDPYCVACFGELFAPKCSSCKRPVIGLGGGK 235  
Db 207 AMGKOWHVEHFVCAKCEKFLGHRHYERKGLAYCETHYNQLFGDVCFHCRNVI-----EGD 262  
QY 236 YVSFDRHHNHCFCARCSTL-VGQGFVDPDQVLCQGC 275  
Db 263 VVSALNKAMCVNCFACSTCNTLTLANKFEVDMKPVCKKC 303

## RESULT 7

US-08-889-402-1  
; Sequence 1, Application US/08889402  
; Patent No. 5811288  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION  
; TITLE OF INVENTION: PROTEIN PAXILLIN  
; NUMBER OF SEQUENCES: 7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,402  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 178334/1996  
; FILING DATE: 08-JUL-1996  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 591 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; CELL TYPE: Monoblastic cell and placenta  
; CELL LINE: U937  
US-08-889-402-1

Query Match 18.6%; Score 304.5; DB 2; Length 591;  
Best Local Similarity 27.5%; Pred. No. 9e-21;  
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;  
QY 40 CAECQQLIGHDSRELFYEDRHFEHCGFCRCRCQSRSLADEPFTQDSELLCNDYCYSAFSS 99  
Db 358 CGACKKPPIA--GOVVTAMGKTWHPHFVCHTCHQEEIGSRNFFRDGPCEKDYHNLFS 415  
QY 100 QCSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRKSFVDPKGAHYVPCYENKFA 159  
Db 416 RCYICNGPIL--DKVVVTALDRTWHPHFVCAQCGAFFGPGHEKDGKAYCRKDYDMEA 473  
QY 160 PRCARCKSLTGGGVTYRDPQPHRECLVCTGCTPLAGQOFTSRDEDDPYCVACFGELFAP 219  
Db 474 PKGGCARAILENYISALNTLWHPFCVCRECFPFVNGSFFFDHGPQYCEVHYHRRGS 533  
QY 220 KCSSCKRPVIGGGKYVSFEDRHHNHCFCARCSTSLVGQGFVDPDQVLCQGC 275  
Db 534 LCSGCKQKPIIT---GRCITAMAKKFFHFEVCAFLKQLNKGTFKQNDKPYCQNC 585

## RESULT 8

US-08-889-402-2  
; Sequence 2, Application US/08889402  
; Patent No. 5811288  
; GENERAL INFORMATION:

; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION  
; TITLE OF INVENTION: PROTEIN PAXILLIN  
; NUMBER OF SEQUENCES: 7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,402  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 178334/1996  
; FILING DATE: 08-JUL-1996  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 605 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; CELL TYPE: Monoblastic cell and placenta  
; CELL LINE: U937  
US-08-889-402-2  
Query Match 18.6%; Score 304.5; DB 2; Length 605;  
Best Local Similarity 27.5%; Pred. No. 9.2e-21;  
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;  
QY 40 CAECQQLIGHDSRELFYEDRHFEHCGFCRCRCQSRSLADEPFTQDSELLCNDYCYSAFSS 99  
Db 372 CGACKKPPIA--GOVVTAMGKTWHPHFVCHTCHQEEIGSRNFFRDGPCEKDYHNLFS 429  
QY 100 QCSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRKSFVDPKGAHYVPCYENKFA 159  
Db 430 RCYICNGPIL--DKVVVTALDRTWHPHFVCAQCGAFFGPGHEKDGKAYCRKDYDMEA 487  
QY 160 PRCARCKSLTGGGVTYRDPQPHRECLVCTGCTPLAGQOFTSRDEDDPYCVACFGELFAP 219  
Db 488 PKGGCARAILENYISALNTLWHPFCVCRECFPFVNGSFFFDHGPQYCEVHYHRRGS 547  
QY 220 KCSSCKRPVIGGGKYVSFEDRHHNHCFCARCSTSLVGQGFVDPDQVLCQGC 275  
Db 548 LCSGCKQKPIIT---GRCITAMAKKFFHFEVCAFLKQLNKGTFKQNDKPYCQNC 599

## RESULT 9

US-09-008-465-1  
; Sequence 1, Application US/09008465  
; Patent No. 6174702  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN PINCH PROTEIN HOMOLOG  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/008,465  
; FILING DATE:



```
; NAME: Wetherell, Jr., Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-3204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-166-316-2

Query Match
Best Local Similarity 17.9%; Score 293; DB 1; Length 454;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY 101 CSAGETVMPGSRKLEYGGQTHWEHCFSLGCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160
Db 279 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKVLEEGGFFFEKGAIFCPCPYDVRYP 336
QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGEFAPK 220
Db 337 SCACRKKKITGEIMHAKMTWVHCFTCAACKTPIRNRAFYMEEGVPYCERDYERKMFGTK 396
QY 221 CSSCKRPIVGLGGKYYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLQ 273
Db 397 CHCDDFKID--AGDRLEALGFSWHDTCFVCAICQINLEKTFYSKKDRPLCK 447

RESULT 12
US-09-124-238A-10
; Sequence 10, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/124,238A
; CURRENT FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-124-238A-10

Query Match
Best Local Similarity 17.9%; Score 293; DB 4; Length 457;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY 101 CSAGETVMPGSRKLEYGGQTHWEHCFSLGCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160
Db 282 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKVLEEGGFFFEKGAIFCPCPYDVRYP 339
QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGEFAPK 220
Db 340 SCACRKKKITGEIMHAKMTWVHCFTCAACKTPIRNRAFYMEEGVPYCERDYERKMFGTK 399
QY 221 CSSCKRPIVGLGGKYYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLQ 273
Db 400 CHCDDFKID--AGDRLEALGFSWHDTCFVCAICQINLEKTFYSKKDRPLCK 450

RESULT 13
US-09-124-238A-10
; Sequence 10, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/124,238A
; CURRENT FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-124-238A-10

Query Match
Best Local Similarity 17.9%; Score 293; DB 4; Length 457;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY 101 CSAGETVMPGSRKLEYGGQTHWEHCFSLGCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160
Db 282 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKVLEEGGFFFEKGAIFCPCPYDVRYP 339
QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGEFAPK 220
Db 340 SCACRKKKITGEIMHAKMTWVHCFTCAACKTPIRNRAFYMEEGVPYCERDYERKMFGTK 399
QY 221 CSSCKRPIVGLGGKYYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLQ 273
Db 400 CHCDDFKID--AGDRLEALGFSWHDTCFVCAICQINLEKTFYSKKDRPLCK 450
```

```
US-09-721-975-10
; Sequence 10, Application US/09721975
; Patent No. 6444803
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/721,975
; CURRENT FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-721-975-10

Query Match
Best Local Similarity 17.9%; Score 293; DB 4; Length 457;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY 101 CSAGETVMPGSRKLEYGGQTHWEHCFSLGCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160
Db 282 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKVLEEGGFFFEKGAIFCPCPYDVRYP 339
QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGEFAPK 220
Db 340 SCACRKKKITGEIMHAKMTWVHCFTCAACKTPIRNRAFYMEEGVPYCERDYERKMFGTK 399
QY 221 CSSCKRPIVGLGGKYYVSFEDRHHNCFSCARCSTSLVGQGFVPDGDQVLQ 273
Db 400 CHCDDFKID--AGDRLEALGFSWHDTCFVCAICQINLEKTFYSKKDRPLCK 450

RESULT 14
US-09-986-621-10
; Sequence 10, Application US/09986621
; Patent No. 6521750
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6521750el Bone Mineralization Proteins, DNA, Vectors,
; FILE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/986,621
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-621-10

Query Match
Best Local Similarity 17.9%; Score 293; DB 4; Length 457;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY 101 CSAGETVMPGSRKLEYGGQTHWEHCFSLGCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160
Db 282 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKVLEEGGFFFEKGAIFCPCPYDVRYP 339
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 20:00:05 ; Search time 57 Seconds

(without alignments)  
648.233 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSSEFCAKCNESLYGRKYI.....QGFPDGDQVLCGCSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 497079 seqs, 131961718 residues

Total number of hits satisfying chosen parameters: 497079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	716	43.8	245	11	US-09-764-891-4041
2	314	19.2	232	15	US-10-023-282-667
3	314	19.2	334	15	US-10-023-282-672
4	313	19.1	386	10	US-09-211-424-2
5	306.5	18.7	314	9	US-09-880-705-2
6	306.5	18.7	314	15	US-10-092-0668-3
7	301	18.4	266	15	US-10-043-487-245
8	301	18.4	596	15	US-10-205-823-222
9	297.5	18.2	341	15	US-10-092-0668-1
10	293	17.9	457	11	US-09-986-625-10
11	293	17.9	498	9	US-09-925-302-475
12	284	17.4	457	11	US-09-986-625-1
13	283	17.3	835	15	US-10-273-680-7
14	275	16.8	831	15	US-10-273-680-2
15	251.5	15.4	300	15	US-10-102-806-684

Sequence 12, Appl  
Sequence 13, Appl  
Sequence 2, Appl  
Sequence 11, Appl  
Sequence 1, Appl  
Sequence 3, Appl  
Sequence 4, Appl  
Sequence 8, Appl  
Sequence 1, Appl  
Sequence 9, Appl  
Sequence 54, Appl  
Sequence 5, Appl  
Sequence 7, Appl  
Sequence 14, Appl  
Sequence 6, Appl  
Sequence 20, Appl  
Sequence 49, Appl  
Sequence 36148, A  
Sequence 143, App  
Sequence 142, App  
Sequence 233, App  
Sequence 234, App  
Sequence 5, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 10, Appl  
Sequence 12, Appl  
Sequence 2, Appl

US-09-779-307-12  
US-09-779-307-13  
US-09-779-307-2  
US-09-779-307-11  
US-10-357-627-1  
US-10-357-627-3  
US-10-147-026-4  
US-09-773-926-17  
US-10-126-099-8  
US-10-126-099-1  
US-10-126-099-9  
US-09-814-122-54  
US-09-938-275-5  
US-10-126-099-7  
US-09-932-367A-14  
US-09-932-367A-6  
US-09-932-367A-20  
US-09-862-027-49  
US-09-862-027-41  
US-09-864-761-36148  
US-09-771-161A-143  
US-09-771-161A-142  
US-09-771-161A-233  
US-09-932-367A-5  
US-09-932-367A-2  
US-09-932-367A-4  
US-09-932-367A-10  
US-09-932-367A-12  
US-10-180-903-2

16 232.5 14.2 547 10  
17 232.5 14.2 547 10  
18 232 14.2 538 10  
19 232 14.2 538 10  
20 220 13.5 421 12  
21 220 13.5 421 12  
22 214.5 13.1 663 15  
23 200 12.2 194 9  
24 200 12.2 194 14  
25 196.5 12.0 193 14  
26 195.5 12.0 193 14  
27 190 11.6 159 9  
28 179.5 11.0 3075 10  
29 178 10.9 193 14  
30 164 10.0 401 11  
31 163.5 10.0 398 11  
32 163 10.0 367 11  
33 162 9.9 615 10  
34 161 9.8 638 10  
35 160.5 9.8 536 9  
36 158.5 9.7 305 10  
37 158.5 9.7 409 10  
38 158.5 9.7 647 10  
39 158.5 9.7 647 10  
40 157.5 9.6 394 11  
41 156.5 9.6 383 11  
42 156.5 9.6 395 11  
43 155.5 9.5 397 11  
44 155.5 9.5 402 11  
45 155.5 9.5 1548 15

#### ALIGNMENTS

#### RESULT 1

US-09-764-891-4041  
; Sequence 4041, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764, 891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 4041  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (86)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (89)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (99)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (116)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (147)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (172)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-4041



RESULT 3  
US-10-023-282-672  
; Sequence 672, Application US/10023282  
; Publication No. US20030092893A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 207 Human Secreted Proteins  
; FILE REFERENCE: P2007P1  
; CURRENT APPLICATION NUMBER: US/10/023,282  
; CURRENT FILING DATE: 2001-12-20  
; EARLIER APPLICATION NUMBER: 09/205,258  
; EARLIER FILING DATE: 1998-12-04  
; EARLIER APPLICATION NUMBER: PCT/US98/11422  
; EARLIER FILING DATE: 1998-06-04  
; EARLIER APPLICATION NUMBER: 60/048,885  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,891  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,971  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,964  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,882  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,899  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,893  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883

; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 60/092,921  
; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 672  
; LENGTH: 334  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (8)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (145)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-023-282-672  
  
Query Match 19.2%; Score 314; DB 15; Length 334;  
Best Local Similarity 32.8%; Pred. No. 1.4e-20;  
Matches 62; Conservative 26; Mismatches 83; Indels 18; Gaps 4;  
  
QY 101 CSACGETVMPGSRKLEYGGTWHCHFLCSCGCEQPIGSGFVDPDKAHYCVPCYENKFP 160  
DB 100 CASCQKPI--AGKVIHALGQSWHPEHFVCTHCKEEIGSSPFFERSGLXCYCPNDYHQLFSP 157  
QY 161 RCARCSKTLTQGGVYRDQPHRECLVCTGCOTPLAGQQFTSRDEDPYCVACFGELFAPK 220  
DB 158 RCAYCAAPILDKVLTAQTWHPHEHFFCSHCGEFGAEGFHEKDKPKYCKRDKFLAMFSPK 217  
QY 221 CSSCKRPYVGLGGKYVSFEDRHHHNCFCARCSTSL-VGGGFVDPGDO----- 269  
DB 218 CGGCNRPVL---ENYLSAMDTVMHPECFVCGDCFTSFSTGTSFFELDGRPFCELHYHRR 273  
QY 270 -VLCQGCSCQ 277  
DB 274 GTLCHGCGQ 282  
  
RESULT 4  
US-09-211-424-2  
; Sequence 2, Application US/09211424A  
; Patent No. US20020177231A1  
; GENERAL INFORMATION:  
; APPLICANT: Staunton et al.  
; TITLE OF INVENTION: Leupaxin Materials and Methods  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/09/211,424A  
; CURRENT FILING DATE: 1998-12-15  
; NUMBER OF SEQ ID NOS: 38  
; SEQ ID NO 2  
; LENGTH: 386  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-211-424-2



Db	90	CHCMQVTR--GPFVLV	LGKSWHEEENCAHCN	TNTMAYIGFVEK	GALYCEJCYEFAP	147
QY	161	RCARSKITL	TOGGTYRDPQ	WHRECLVCTG	QTPLAQQT	220
Db	148	EGRCOR	KILGEVINAL	QKTHVSCF	VCAKGRPI	207
QY	221	CSSCRPI	VLGGKYYVS	FEDRHH	NCFSCAR	273
Db	208	CHGCFE	PIE--AGDMF	LEALGYT	WHDDTCF	258

## RESULT 8

```

US-10-205-823-222
; Sequence 222, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersche, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumel
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 222
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-222

```

## RESULT 9

US-10-092-066B-1  
; Sequence 1, Application US/10092066B  
; Publication No. US20030104472A1

: GENERAL INFORMATION:

```

1  GENERAL INFORMATION:
2  APPLICANT: Lal, Preeti G.
3  APPLICANT: Guegler, Karl J.
4  APPLICANT: Corley, Neil C.
5  TITLE OF INVENTION: ANTIBODY SPECIFICALLY BINDING HUMAN PINCH PROTEIN HOMOLOG
6  FILE REFERENCE: PF-0460-2C1P
7  CURRENT APPLICATION NUMBER: US/10/092,066B
8  CURRENT FILING DATE: 2002-03-04
9  PRIOR APPLICATION NUMBER: 09/528,959
10 PRIOR FILING DATE: 2000-03-20
11 PRIOR APPLICATION NUMBER: 09/008,465
12 PRIOR FILING DATE: 1998-01-16
13 NUMBER OF SEQ ID NOS: 3
14 SOFTWARE: PERL Program
15 SEQ ID NO 1
16 LENGTH: 341
17 TYPE: PRT
18 ORGANISM: Homo sapiens
19 FEATURE:
20 NAME/KEY: misc_feature
21 OTHER INFORMATION: Incyte ID No. US20030104472A1 3540806CD1
22 US-10-092-066B-1

```

## RESULT 10

```

RESOLVING: 10
US-09-986-625-10
; Sequence 10, Application US/09986625
; Publication No. US20030125248A1
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. US20030125248A1el
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148-0115
; CURRENT APPLICATION NUMBER: US/09/986,625
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 10
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-625-10

```

```

Query Match          17.9%; Score 293; DB 11; Length 457;
Best Local Similarity 32.4%; Pred. No. 1.6e-18;
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;

QY      101 CSACGETVMGSRKLEYGGQWTWHEHCFLCSCCEOPLGSRSPVDPKGAHYCVPCYENKAP 160
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      282 CHQCHKVIR--GRYLVALGHAYHPPEEFVCSQCGKVLBEGGFFEEKGAIFCPCPYDVRYP 339
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY      161 RCARCSKTLTQGGVTYRDQPHRRCCLVCTGCOTPLAQOQFTRDEDPYCVACFGELFAPK 220
      || : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      340 SCARCKKKKITGEINHALKMTWHVHCFTCAACKTPIRNRAFYMEGVYPYCERDYKEMFGTK 399
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY      221 CSSCKRPVIGLGGGKYVSFEDRHHWHNCFSCARCSSTSLVGQGFVDPDQGVLCQ 273
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db      400 CHGCDFKID--AGDRFLEALGFSHWDTCFVCAICQINLEGKTFYSKKDRPLCK 450
      | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

RESULT 11
US-09-925-302-475
; Sequence 475, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 475
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (67)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (119)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (120)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-475

```

```

; Publication No. US20030125248A1
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. US20030125248A1el Bone Mineralization Proteins, DNA, Vecto
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/986,625
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 09/124,238
; PRIOR FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ_ID NO 1
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-986-625-1

Query Match          17.4%; Score 284; DB 11; Length 457;
Best Local Similarity 31.8%; Pred.No.1e-17;
Matches 55; Conservative 27; Mismatches 87; Indels 4; Gaps 2;

QY      101 CSACGETVMPGSRKLEYGGQTWHEHCFLCSCGEQPLGSRFSVPDGAHYCVPCYENKFAP 160
        | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db       282 CHQCHKIIR--GRYLVALGHAYHPPEFVCSQCCKVLEEGGFEEKGAIFCPSCYDVRYAP 339

QY      161 RCARCSKTLTQGVTYRDQPWHRECLVCTGCOTPLAGOOFTSRDEDPYCVCAGFELFAPK 220
        |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db       340 SCACKKKITGETINHALKMTWTPCTCAACKTKPIRNRAFYMEEGAPICERYDKMFGTK 399

QY      221 CSSCKRPIVLGGGKGVYSFEDRRHHNHCFSARCSTSLVGQQGFVPDGDOVLQO 273
        |||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db       400 CRCGRDKID--ADGRFLAALGFSWHDTCFWCAICNTINLEGKTFYSKKDKPLCK 450

```

## RESULT 12

US-09-986-625-1

00 02 00 025 1  
; Sequence 1, Application US/09986625

Db 98 -----QMSAQRKKEALGRGNIKMLSRVHMATCEKCEKINGGEVAIFVSRAGPGVC 150  
QY 122 WHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRD-QP 180  
Db 151 WHPSCFVCTCNELLDVLIYFYQDGKHGRHHAELLKPCRSACDELIIFADECTEAEGRH 210  
QY 181 WHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRD-QP 240  
Db 211 WHMHFFCCYBCTETVLGQRYIMKDRPFCCGCGPESHVAEYCEGGEHI-GVDHAQ-MTYD 268  
QY 241 DRWHH--HNCFCARCSTSLVGGQFVDPGDQV-LCOGCS 276  
Db 269 GQHWATETCFSCAQCKVSLGCPFLPKGRYICCKACS 307

RESULT 14

US-10-273-680-2  
; Sequence 2, Application US/10273680  
; Publication No. US2003009996A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Tomlinson, James E.  
; APPLICANT: Komuves, Laszlo G.  
; TITLE OF INVENTION: VELD2, A VASCULAR ENDOTHELIAL CELL  
; TITLE OF INVENTION: SPECIFIC AND LIM DOMAIN CONTAINING MOLECULE AND USES  
; FILE REFERENCE: MPI01-317PIRNM  
; CURRENT APPLICATION NUMBER: US/10/273,680  
; CURRENT FILING DATE: 2002-10-18  
; PRIOR APPLICATION NUMBER: US 60/329,756  
; PRIOR FILING DATE: 2001-10-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 831  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
US-10-273-680-2

Query Match 16.8%; Score 275; DB 15; Length 831;  
Best Local Similarity 27.0%; Pred. No. 1.2e-16;  
Matches 76; Conservative 43; Mismatches 106; Indels 56; Gaps 13;  
QY 28 CVCYDNTFANTCAEQQLIGHDSRELFYEDRHFEGCFRCRCORSLADEPFTCODSEL 87  
Db 51 CLPEERKVPVNSPGE-----KHRIKQLLYQ-LPPHNEVRYC---QSLSEE-----EKKEL 97  
QY 88 LCNDCYCSAFSSQ-----CSAC-----GETVMPGSRKLEYGG 119  
Db 98 -----QVFSQRKKEALGRGTIKLLSRVHMATCEKCEKINGGEVAIFVSRAGPGVC 148  
QY 120 QTWHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRD- 178  
Db 149 VCMHPSCFVCTCNELLDVLIYFYQDGKHGRHHAELLKPCRSACDELIIFADECTEAE 208  
QY 179 QWHEHCFCLSCGEQPLGSRSPVDPKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRD- 238  
Db 209 RWHMHFFCCYBCTETVLGQRYIMKDRPFCCGCGPESHVAEYCEGGEHI-GVDHAQ-MT 266  
QY 239 FEDRWHH--HNCFCARCSTSLVGGQFVDPGDQV-LCOGCS 276  
Db 267 YDQHWATEACFSCAQCKVSLGCPFLPKGRYICCKACS 307

RESULT 15

US-10-102-806-684  
; Sequence 684, Application US/10102806  
; Publication No. US20030054421A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA103PIC1

; CURRENT APPLICATION NUMBER: US/10/102,806  
; CURRENT FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/925,298  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05881  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 846  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 684  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-102-806-684  
Query Match 15.4%; Score 251.5; DB 15; Length 300;  
Best Local Similarity 28.0%; Pred. No. 5.5e-15;  
Matches 61; Conservative 25; Mismatches 97; Indels 35; Gaps 5;  
QY 69 CRCORSIADBPFTCQDSSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGTWHHCFL 128  
Db 106 CGCQTLRPM-----STPGGAAAVRALGQLFHIACFT 139  
QY 129 CSCEQPLGSRSPVDPKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRDQWHEHCLVC 188  
Db 140 CHCAQOLQOQQFYSLGAPYCEGYDTL-EKNTCGEPIITDMLRATGKAYHPHCFTC 198  
QY 189 TGCQTPLAGQOF-TSRDEDDYCVACFGELEFAPKCSSCKRPV---GLGGKYYVSFEDRHW 244  
Db 199 VVCARPLEGTSFTVDQANRPHCVDPYHKQYAPRCVSEPIPEPGRDETFRVVALDKNF 258  
QY 245 HNCFCARCSTSLV---GQGFVDPGDQVLCOGCSQA 278  
Db 259 HMKYKCEDCGKPLSTEADNNGCFPLDGHVLCRCKHTA 296

Search completed: August 22, 2003, 20:09:25  
Job time : 59 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:57:47 ; Search time 384 Seconds  
(without alignments)  
634.636 Million cell updates/sec

Title: US-09-786-135-1  
Perfect score: 1635  
Sequence: 1 MSESFDKCNESLYGRKYI.....QGFPDGDVLCQGCQAGP 280

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues  
Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1:	/cgn2_6/ptodata/1/paa/US06_COMB.pap.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pap.*
3:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
5:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
6:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
7:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
8:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
9:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
10:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
11:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
12:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
13:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
14:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
15:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
16:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
17:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
18:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
19:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
20:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
21:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
22:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
23:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
24:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
25:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
26:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
27:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
28:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
29:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
30:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*
31:	/cgn2_6/ptodata/1/paa/US08_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1635	100.0	280	22	US-09-786-135-1 Sequence 1, Appl1
2	1635	100.0	280	31	US-60-453-050-9996 Sequence 9996, Ap

3	1635	100.0	280	31	US-60-453-135-9996 Sequence 9996, Ap
4	1635	100.0	280	31	US-60-453-135-9996 Sequence 9996, Ap
5	1635	100.0	280	24	US-09-949-016-9099 Sequence 9099, Ap
6	1625	99.4	280	22	US-09-791-537-141395 Sequence 141395, Ap
7	1615	98.8	280	22	US-09-791-537-141479 Sequence 141479, Ap
8	1595	97.6	280	21	US-09-724-676-62747 Sequence 62747, A
9	1578.5	96.5	280	21	US-09-724-676A-62747 Sequence 62747, A
10	1569	96.0	280	22	US-09-791-537-141433 Sequence 141433, Ap
11	1563	95.6	280	22	US-09-791-537-141432 Sequence 141432, Ap
12	998	61.0	279	22	US-09-791-537-87691 Sequence 87691, A
13	991	60.6	279	22	US-09-791-537-2014 Sequence 2014, Ap
14	974	59.6	279	21	US-09-724-676-62522 Sequence 62522, A
15	974	59.6	279	21	US-09-724-676-62527 Sequence 62527, A
16	974	59.6	279	21	US-09-724-676-62538 Sequence 62538, A
17	974	59.6	279	21	US-09-724-676A-62522 Sequence 62522, A
18	974	59.6	279	21	US-09-724-676A-62527 Sequence 62527, A
19	974	59.6	279	21	US-09-724-676A-62538 Sequence 62538, A
20	974	59.6	279	22	US-09-791-537-128113 Sequence 128113, Ap
21	974	59.6	279	31	US-60-389-987-169 Sequence 169, App
22	974	59.6	279	31	US-60-412-418-169 Sequence 169, App
23	974	59.6	279	31	US-60-452-680-21669 Sequence 21669, A
24	974	59.6	279	31	US-60-452-680-21670 Sequence 21670, A
25	974	59.6	279	31	US-60-452-680-21670 Sequence 21670, A
26	968	59.2	327	28	US-10-206-021-538 Sequence 100836, Ap
27	951.5	58.2	304	21	US-09-724-676-62520 Sequence 62520, A
28	951.5	58.2	304	21	US-09-724-676-62521 Sequence 62521, A
29	951.5	58.2	304	21	US-09-724-676-62525 Sequence 62525, A
30	951.5	58.2	304	21	US-09-724-676-62526 Sequence 62526, A
31	951.5	58.2	304	21	US-09-724-676-62530 Sequence 62530, A
32	951.5	58.2	304	21	US-09-724-676-62531 Sequence 62531, A
33	951.5	58.2	304	21	US-09-724-676A-62520 Sequence 62520, A
34	951.5	58.2	304	21	US-09-724-676A-62521 Sequence 62521, A
35	951.5	58.2	304	21	US-09-724-676A-62525 Sequence 62525, A
36	951.5	58.2	304	21	US-09-724-676A-62526 Sequence 62526, A
37	951.5	58.2	304	21	US-09-724-676A-62530 Sequence 62530, A
38	951.5	58.2	304	21	US-09-724-676A-62531 Sequence 62531, A
39	951.5	58.2	304	22	US-09-791-537-143019 Sequence 143019, Ap
40	866	53.0	284	19	US-09-522-689A-2 Sequence 2, Appl1
41	866	53.0	284	19	US-09-522-689A-2 Sequence 2, Appl1
42	866	53.0	284	22	US-09-791-537-141508 Sequence 141508, Ap
43	865	52.9	284	22	US-09-791-537-141507 Sequence 141507, Ap
44	865	52.9	284	22	US-09-791-537-141507 Sequence 141507, Ap
45	865	52.9	290	24	US-09-949-016-11587 Sequence 11587, A

ALIGNMENTS

RESULT 1  
US-09-786-135-1  
; Sequence 1, Application US/09786135  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: PATTERSON, Chandra  
; TITLE OF INVENTION: LIM DOMAIN PROTEIN HOMOLOGY  
; FILE REFERENCE: PF-0585 PCT  
; CURRENT APPLICATION NUMBER: US/09/786,135  
; PRIOR FILING DATE: 1998-09-03; 1998-09-03  
; PRIOR FILING DATE: 1998-09-03; 1998-09-03  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No: 1925017CDI  
US-09-786-135-1



Db 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
QY 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
Db 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 280  
Db 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 280

RESULT 5

US-09-949-016-9099  
; Sequence 9099.. Application US/09949016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9099  
; LENGTH: 287  
; TYPE: PRP  
; ORGANISM: Human  
US-09-949-016-9099

Query Match 100.0%; Score 1635; DB 24; Length 287;  
Best Local Similarity 100.0%; Pred. No. 2.1e-121;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 60  
Db 8 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 67  
QY 61 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120  
Db 68 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 127  
QY 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
Db 128 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 187  
QY 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
Db 188 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 247  
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 280  
Db 248 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 287

RESULT 6

US-09-791-537-141395  
; Sequence 141395.. Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 141395  
; LENGTH: 280  
; TYPE: PRP  
; ORGANISM: Mus musculus  
US-09-791-537-141395

Query Match 99.4%; Score 1625; DB 22; Length 280;  
Best Local Similarity 99.3%; Pred. No. 1.3e-120;  
Matches 278; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 60  
Db 1 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 60  
QY 61 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120  
Db 61 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
Db 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
QY 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
Db 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
QY 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 280  
Db 241 DRWHHNCFCARCSSTSLVGQGFVDPDQVLCGCSQAGP 280

RESULT 7

US-09-791-537-141479  
; Sequence 141479.. Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 141479  
; LENGTH: 280  
; TYPE: PRP  
; ORGANISM: Homo sapiens  
US-09-791-537-141479

Query Match 98.8%; Score 1615; DB 22; Length 280;  
Best Local Similarity 98.9%; Pred. No. 8.1e-120;  
Matches 277; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 60  
Db 1 MSSEFDCAKCNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFLYEDRH 60  
QY 61 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120  
Db 61 FHGECFRCRCQSLADEPFTCDSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
Db 121 TWHEHCFCLSCGCEQPLGSRSFVDPKGAHYCPVCYENKFAFPCARCSKTLTQGGVYTRDQP 180  
QY 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240  
Db 181 WHRECLVCTGCTPLAGQQTSDRDEDPYCACFGELFAPKCSCKRPIVGLGGKYYVSFE 240

QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280  
|||||  
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

## RESULT 8

US-09-724-676-62747  
; Sequence 62747, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Comugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Comugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 62747  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-62747

Query Match 97.6%; Score 1595; DB 21; Length 280;  
Best Local Similarity 97.1%; Pred. No. 3.1e-118;  
Matches 272; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
|||||  
Db 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
QY 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
|||||  
Db 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
|||||  
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 240  
|||||  
Db 181 WHPKCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 240  
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280  
|||||  
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

## RESULT 9

US-09-724-676A-62747  
; Sequence 62747, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Comugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Comugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 62747  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-62747

Query Match 97.6%; Score 1595; DB 21; Length 280;  
Best Local Similarity 97.1%; Pred. No. 3.1e-118;  
Matches 272; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
|||||  
Db 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
QY 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120

Db 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 240  
Db 181 WHPKCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 240  
QY 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280  
Db 241 DRWHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280

## RESULT 10

US-09-791-537-143101  
; Sequence 143101, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biocomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 143101  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-143101

Query Match 96.5%; Score 1578.5; DB 22; Length 289;  
Best Local Similarity 93.4%; Pred. No. 6.6e-117;  
Matches 270; Conservative 6; Mismatches 4; Indels 9; Gaps 1;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
|||||  
Db 1 MSESFDCAKCNESLYGRKYIQTDSGPYCPYDNTFANTCAECQQLIGHDSRELFEYDRH 60  
QY 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
|||||  
Db 61 FHEGFCRCRCQSLADEPFTCQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
Db 121 TWHEHCFCLSCGCEQPLGSRFVDPDKGAHYCPYENKFAPRCARCSKTLTQGGVTVYRDP 180  
QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 231  
Db 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKYVSFE 231  
QY 232 GGGKYVSFEEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 280  
Db 241 GGGKYVSFEEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQGCQAGP 289

## RESULT 11

US-09-791-537-141433  
; Sequence 141433, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biocomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 141433  
; LENGTH: 280  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-141433

Query Match 96.0%; Score 1569; DB 22; Length 280;  
Best Local Similarity 95.7%; Pred. No. 3.6e-116;  
Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60  
Db 1 MSESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60  
QY 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
Db 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
Db 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
QY 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240  
Db 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240  
QY 241 DRHWHNCFSCARCSSTSLVGQGFVDPDQVLCQGCQSQAGP 280  
Db 241 DRHWHNCFSCARCSSTSLVGQGFVDPDQVLCQGCQSQAGP 280

RESULT 12

US-09-791-537-141432  
; Sequence 141432, Application US/09791537

; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 141432  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-141432

Query Match 95.6%; Score 1563; DB 22; Length 280;  
Best Local Similarity 95.4%; Pred. No. 1.1e-115;  
Matches 267; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60  
Db 1 MSESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60  
QY 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
Db 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
Db 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
QY 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240  
Db 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240

QY 241 DRHWHNCFSCARCSSTSLVGQGFVDPDQVLCQGCQSQAGP 280  
Db 241 DRHWHNCFSCARCSSTSLVGQGFVDPDQVLCQGCQSQAGP 280

RESULT 13

US-09-791-537-87691  
; Sequence 87691, Application US/09791537

; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 87691  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-791-537-87691

Query Match 61.0%; Score 998; DB 22; Length 279;  
Best Local Similarity 55.6%; Pred. No. 8.1e-71;  
Matches 154; Conservative 55; Mismatches 68; Indels 0; Gaps 0;

QY 1 MSESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRELFEYEDRH 60  
Db 1 MTERFDCCHNCSLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLSYKDRH 60  
QY 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
Db 61 FHGGFCRCRCORSRLADEPFTQDSLELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
QY 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
Db 121 TWHEHCFCLSGCEQPLGSRSFYDPAKGAHYCVPCYENKFAAPRCARCSKTILTOGGVTVRDQP 180  
QY 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240  
Db 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYSVSE 240  
QY 241 DRHWHNCFSCARCSSTSLVGQGFVDPDQVLCQGCQSQ 277  
Db 241 ERQWHDNCFNCKKCSLSLVGRGFLTERDDILCPDCGK 277

RESULT 14

US-09-791-537-2014  
; Sequence 2014, Application US/09791537

; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2014  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-2014

Query Match 60.6%; Score 991; DB 22; Length 279;  
Best Local Similarity 54.9%; Pred. No. 2.9e-70;  
Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;

```
QY 1 MSESFDCAKNESLYGRKYIQTDSGPCYPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60
Db 1 MTERFDCHHCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLISKDRH 60
QY 61 FHEGCFRCRCQRSLADEPFTCQDSSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 WHEGCFHCSRGSSLVDPKFAAKEQELLCDCYSNEYSSKQCECKKTIMPGRMEYKGS 120
QY 121 TWHEHCFCLSCGEQPLGSRGFVPDKGAHYCPVCYENKFAPCARCSKTLTGGVYTRDOP 180
Db 121 SWHETCFQCRCQPIGTGKSFIPKFNQNFVPCYKQYALQCVQCKPITTTGGVYTRDOP 180
QY 181 WHRECLVTCQQTPLAGQQTFRDEDPYCIVACFGELFAPKSCSKRPITVGLGGKYYSFE 240
Db 181 WHKECFVCTACKKQLSGORFTARDEFFVCLTCFCDLYAKKACAGCTNPISGLGGTKYISFE 240
QY 241 DRWHHNCFCARCSSTLVGGVFDGQVLCQGCSCQ 277
Db 241 EROWHNDCFNCKGCSLVGRGFLTERDDILCPDCGK 277
```

## RESULT 15

```
US-09-724-676-62522
; Sequence 62522, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62522
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-62522
```

```
Query Match 59.6%; Score 974; DB 21; Length 279;
Best Local Similarity 53.8%; Pred. No. 6.5e-69;
Matches 149; Conservative 59; Mismatches 69; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIQTDSGPCYPCYDNTFANTCAEQOOLIGHDSRELFYEDRH 60
Db 1 MTERFDCHHCNESLYGKYLKEENPHCVACFEELYANTCECGTPIGCDCKDLISKDRH 60
QY 61 FHEGCFRCRCQRSLADEPFTCQDSSELLCNDYCYSFSSQCSAGETVMPGSRKLEYGGQ 120
Db 61 WHEGCFHCSRGSSLVDPKFAAKEQELLCDCYSNEYSSKQCECKKTIMPGRMEYKGS 120
QY 121 TWHEHCFCLSCGEQPLGSRGFVPDKGAHYCPVCYENKFAPCARCSKTLTGGVYTRDOP 180
Db 121 SWHETCFQCRCQPIGTGKSFIPKFNQNFVPCYKQYALQCVQCKPITTTGGVYTRDOP 180
QY 181 WHRECLVTCQQTPLAGQQTFRDEDPYCIVACFGELFAPKSCSKRPITVGLGGKYYSFE 240
Db 181 WHKECFVCTACKKQLSGORFTARDEFFVCLTCFCDLYAKKACAGCTNPISGLGGTKYISFE 240
QY 241 DRWHHNCFCARCSSTLVGGVFDGQVLCQGCSCQ 277
Db 241 EROWHNDCFNCKGCSLVGRGFLTERDDILCPDCGK 277
```

Search completed: August 22, 2003, 20:07:49  
Job time : 386 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:58:12 : Search time 24 Seconds  
(without alignments)  
313.637 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCAKNESLYGRKYI.....QGFPDQVLCQGCQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 128601 seqs, 26883176 residues

Total number of hits satisfying chosen parameters: 128601

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending\_Patents\_AA\_New.\*
- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pap.\*
  - 2: /cgn2\_6/ptodata/1/paa/US05\_NEW\_COMB.pap.\*
  - 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pap.\*
  - 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pap.\*
  - 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pap.\*
  - 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pap.\*
  - 7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	974	59.6	279	6	US-10-408-765A-169
2	974	59.6	279	7	US-60-490-890-1761
3	437.5	26.8	150	6	US-10-408-765A-210
4	305.5	18.7	264	7	US-60-487-610-1707
5	305.5	18.7	461	6	US-10-293-244-1620
6	305.5	18.7	486	6	US-10-293-244-3586
7	303.5	18.6	734	6	US-10-408-765A-324
8	292.5	17.9	296	6	US-10-374-979-102
9	275	16.8	831	6	US-10-408-765A-2503
10	265	16.2	615	6	US-10-293-244-1901
11	265	16.2	615	6	US-10-408-765A-913
12	262	16.0	539	6	US-10-408-765A-1767
13	260	15.9	212	6	US-10-275-595A-21
14	257.5	15.7	572	6	US-10-374-979-107
15	245.5	15.0	645	6	US-10-293-244-3869
16	228	13.9	431	6	US-10-374-979-90
17	220	13.5	412	7	US-60-490-890-379
18	220	13.5	421	6	US-10-408-765A-1032
19	220	13.5	421	6	US-10-286-897-2375
20	220	13.5	421	6	US-10-258-898A-2375
21	220	13.5	421	7	US-60-490-890-377
22	220	13.5	437	6	US-10-286-897-5947
23	220	13.5	437	6	US-10-258-898A-5947
24	209	12.8	676	6	US-10-374-979-101
25	200	12.2	194	6	US-10-291-172-224
26	196.5	12.0	365	6	US-10-286-897-3084

ALIGNMENTS

RESULT 1

US-10-408-765A-169

; Sequence 169, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 169

; LENGTH: 279

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-408-765A-169

Query Match 59.6%; Score 974; DB 6; Length 279;

Best Local Similarity 53.8%; Pred. No. 4.le-80;

Matches 149; Conservative 59; Mismatches 69; Indels 0; Gaps 0;

Qy 1 MSESFDCAKNESLYGRKYIQTDSGPYCVPCYDNTFANTCAECQOLIGHDSRELFLYEDRH 60

Db 1 MTERFDCCHCNESLFGKKYILRESPYCVVCFETLFTANTCECGKPIGCDCKLSYKDRH 60

Qy 61 FHEGCFCCRCORSLADEPTCQDSLELLCNDYCASFSSQCSAGETVMPGSRKLEYGGQ 120

Db 61 WHEACFHCSCRNLLVDKPFAAKEDQLCTDYSNEYSSKQCECKKTIIMPGRKMEYKGS 120

Qy 121 TWHEHCFCLSCGEOPGLGSRFPDGAHYCPVCYENKFAPCARCSKTLTQGGVTRYDQP 180

Db 121 SWHETCFICHRCCQPIGTFSKSPFKDQNFVPCYECYQHAMQCVCKKPIITGGVTRYDQP 180

Qy 181 WHRECLVCTGQCTPLAQOQFTSRDEDPYCVACFELFAPKCSSCKRPILVGGGKYVSFE 240

Db 181 WHKECFVCTACRKLQSGORFTARDFAVLCNCFDLVAKKACAGCTNPISGLGTYISFE 240

Qy 241 DRWHHNCFCARCSSTLVGQGFVDPDGVLCQGCSCQ 277

Db 241 ERQWHNDFCNCKCSLSLVGRGFLTERDDILCPDCKG 277

## RESULT 2

US-60-490-890-1761  
; Sequence 1761, Application US/60490890  
; GENERAL INFORMATION:  
; APPLICANT: Li, Martha  
; APPLICANT: Rupnow, Brent A.  
; APPLICANT: Webster, Kevin R.  
; APPLICANT: Jackson, Donald  
; APPLICANT: Wong, Tai W.  
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION  
; FILE REFERENCE: D0310 PSP  
; CURRENT APPLICATION NUMBER: US/60/490,890  
; CURRENT FILING DATE: 2003-07-29  
; NUMBER OF SEQ ID NOS: 2779  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1761  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-490-890-1761

Query Match 59.6%; Score 974; DB 7; Length 279;  
Best Local Similarity 53.8%; Pred. No. 4, 1e-80;  
Matches 149; Conservative 59; Mismatches 69; Indels 0; Gaps 0;  
QY 1 MSESFDCAKNESLYGRKIYQTDGSPYCVPCYDNTFANTCAEQQLIGHDSRELFYEDRH 60  
D 1 MTERFDCHHCNESLFGKYLRESPYCVWCFETLFAFNTCECGKPIGCDCKDLKSYKDRH 60  
QY 61 FHEGCFRCRCORSADPFFTCQSELLCNCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
D 61 WHEACFFHCQCRNSLVDPFAAKEDQLLCTDCYSNEYSKQCECKKTIIMPGRKMEYKGS 120  
QY 121 TWHEHCFCLSCGCEPGLSGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGQGVYRQDP 180  
D 121 SWHETCFCHRCQPIGRTKSLPRDNQNFVPCYEKQHAMQCVQCKKPIITGGVYRQDP 180  
QY 191 WHRECLVTCGOTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPVLGGLGGKYVSFE 240  
D 181 WHKECFVCTACRKQLSGQRTARDDFAYCLNCFCDLYAKKAGCTNPISGLGGTKYISFE 240  
QY 241 DRWHHNCFCSCARCSSTLVGGGVDPDGVQVLCQCSQ 277  
D 241 ERQWHDNCFNCKKCSLVGRGFLTERDDILCPDGGK 277

## RESULT 3

US-10-408-765A-210  
; Sequence 210, Application US/10408765A  
; GENERAL INFORMATION:  
; APPLICANT: Ghosh, Soumitra S.  
; APPLICANT: Fahy, Eoin D.  
; APPLICANT: Zhang, Bing  
; APPLICANT: Gibson, Bradford W.  
; APPLICANT: Taylor, Steven W.  
; APPLICANT: Glenn, Gary M.  
; APPLICANT: Warnock, Dale E.  
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
; FILE REFERENCE: 660086.465  
; CURRENT APPLICATION NUMBER: US/10/408,765A  
; CURRENT FILING DATE: 2003-04-04  
; NUMBER OF SEQ ID NOS: 3077  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 210  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-408-765A-210

Query Match 26.8%; Score 437.5; DB 6; Length 150;  
Best Local Similarity 45.3%; Pred. No. 2, 2e-32;  
Matches 68; Conservative 26; Mismatches 55; Indels 1; Gaps 1;

QY 106 ETVMPSRKLEYGGQTWHEHCFCLSCGCEPGLSGSRFVDPKGAHYCVPCYENKFAFPCARC 165  
D 1 KAIIVAGDONVEYKGTVMHKDCFTCSNCKQVIGTGSFFPKGEDFYCVTCHETKLAKHKVC 60  
QY 166 SKTLTQGGVYRDPQWHRRECLVTCGOTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCK 225  
D 61 NKAITSGGITYQDQPHWADCFVVCVTSKLAGQRTAVEDQYCYDCYKKNFVAKKAGCK 120  
QY 226 RPVGLG-GGKYVSPEDRHHHNCFCSCAR 254  
D 121 NPITGFGKSSVVAYEGGSHWHDYCFHCKKC 150

## RESULT 4

US-60-487-610-1707  
; Sequence 1707, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1707  
; LENGTH: 264  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-1707

Query Match 18.7%; Score 305.5; DB 7; Length 264;  
Best Local Similarity 31.4%; Pred. No. 2, 5e-20;  
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;  
QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEPGLSGSRFVDPKGAHYCVPCYENKFA 160  
D 31 CGSCNKP1--AGQVVTALGRAWHPEHFVCGGCGSTALGSSFFEKDGAPPCPCYERFSP 88  
QY 161 RCARCSKTLTGQGVYRDPQWHRRECLVTCGOTPLAGQOFTSRDEDPYCVACFGELFAPK 220  
D 89 RCGFCNQPIRHKMVTALGTHWHPEHFCVSCGEPFGDEGFHREGRPYCRDRDLQLFAPR 148  
QY 221 CSSCKRPVGLGGKYVSFEDRHHHNCFCSCARCSSTLVGGGVDPDGVQVLCQG----- 274  
D 149 CGCGGPIL----DNYISALSALWHPDCVCRECFAPFGSGSFFEHGGRPLCENHFHARR 204  
QY 275 ---CSQAG 279  
D 205 GSLCATCG 212

## RESULT 5

US-10-293-244-1620  
; Sequence 1620, Application US/10293244  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang, Y. Tom et al  
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-029  
; CURRENT APPLICATION NUMBER: US/10/293,244  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: Not Yet Assigned  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 09/728,422  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 09/693,325  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/663,561  
; PRIOR FILING DATE: 2000-09-15



```

; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1620
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1620

Query Match      18.7%; Score 305.5; DB 6; Length 461;
Best Local Similarity 31.4%; Pred. No. 4.6e-20;
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYENKFP 160
Db 228 CGSCNKPI--AGQVVTALGRAWHPHFVCGGCGTALGGSSFFFEKDGAPFCPECYFERFSP 285
QY 161 RCARCSKTLTQGGVTVRDPQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 286 RCQFCNQPIRHKVMVTALGTHWHPEHFCVSCGEPFGDEGHEHREGRPCRRDLQLFAPR 345
QY 221 CSACKRPVGLGGKGVSFEDRHHNHCFCARCSSTSLVGQGFVPDGDQVLCQG----- 274
Db 346 CQCGQGPIL----DNVISALSALWHPDCFCVRCRCFAPFGSGGFFEGHGRPLCENHFHARR 401
QY 275 ---CSQAG 279
Db 402 GSLCATCG 409

RESULT 6
US-10-293-244-3588
; Sequence 3588, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 3588
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-293-244-3588
Query Match      18.7%; Score 305.5; DB 6; Length 461;
Best Local Similarity 31.4%; Pred. No. 4.6e-20;
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYENKFP 160
Db 228 CGSCNKPI--AGQVVTALGRAWHPHFVCGGCGTALGGSSFFFEKDGAPFCPECYFERFSP 285
QY 161 RCARCSKTLTQGGVTVRDPQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 286 RCQFCNQPIRHKVMVTALGTHWHPEHFCVSCGEPFGDEGHEHREGRPCRRDLQLFAPR 345
QY 221 CSACKRPVGLGGKGVSFEDRHHNHCFCARCSSTSLVGQGFVPDGDQVLCQG----- 274
Db 346 CQCGQGPIL----DNVISALSALWHPDCFCVRCRCFAPFGSGGFFEGHGRPLCENHFHARR 401
QY 275 ---CSQAG 279
Db 402 GSLCATCG 409

RESULT 7
US-10-408-765A-324
; Sequence 324, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 324
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-324

Query Match      18.6%; Score 303.5; DB 6; Length 734;
Best Local Similarity 36.3%; Pred. No. 1.1e-19;
Matches 65; Conservative 21; Mismatches 86; Indels 7; Gaps 5;

QY 98 SSQCSACG--ETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYE 155
Db 552 SSRTPLGCHGNVIRGP-FLVAMGRSWHPPEFTCAKTSADVCFVEEQNNVYCERCYE 610
QY 156 NKFAPCARCSKTLTQGGVTVR-DQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFG 214
Db 611 QFFAPLCAKCN-TKMGVFMHALRQTHWTTTCFVCAACKKPGNSLFHMEDEGPYCKDYI 669
QY 215 ELFAPKCSSCKRPVGLGGKGVSFEDRHHNHCFCARCSSTSLVGQGFVPDGDQVLCQ 273
Db 670 NLFSTKCHGCDPVE--AGDKFIEALGHTWHTDTCFICAVCHVNLEQGFYSSKDRPLCK 726

RESULT 8
US-10-374-979-102
; Sequence 102, Application US/10374979
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
```

```

US-10-293-244-3588
Query Match      18.7%; Score 305.5; DB 6; Length 486;
Best Local Similarity 31.4%; Pred. No. 4.9e-20;
Matches 59; Conservative 24; Mismatches 90; Indels 15; Gaps 3;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYENKFP 160
Db 253 CGSCNKPI--AGQVVTALGRAWHPHFVCGGCGTALGGSSFFFEKDGAPFCPECYFERFSP 310
QY 161 RCARCSKTLTQGGVTVRDPQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFGELFAPK 220
Db 311 RCQFCNQPIRHKVMVTALGTHWHPEHFCVSCGEPFGDEGHEHREGRPCRRDLQLFAPR 370
QY 221 CSACKRPVGLGGKGVSFEDRHHNHCFCARCSSTSLVGQGFVPDGDQVLCQG----- 274
Db 371 CQCGQGPIL----DNVISALSALWHPDCFCVRCRCFAPFGSGGFFEGHGRPLCENHFHARR 426
QY 275 ---CSQAG 279
Db 427 GSLCATCG 434

RESULT 7
US-10-408-765A-324
; Sequence 324, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 324
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-324

Query Match      18.6%; Score 303.5; DB 6; Length 734;
Best Local Similarity 36.3%; Pred. No. 1.1e-19;
Matches 65; Conservative 21; Mismatches 86; Indels 7; Gaps 5;

QY 98 SSQCSACG--ETVMPGSRKLEYGGQTWHEHCFCLSCGCEQPLGSRFSFVDPDKGAHYCVPCYE 155
Db 552 SSRTPLGCHGNVIRGP-FLVAMGRSWHPPEFTCAKTSADVCFVEEQNNVYCERCYE 610
QY 156 NKFAPCARCSKTLTQGGVTVR-DQWHRRECLVCTGCQTPLAGQQFTSRDEDPYCVACFG 214
Db 611 QFFAPLCAKCN-TKMGVFMHALRQTHWTTTCFVCAACKKPGNSLFHMEDEGPYCKDYI 669
QY 215 ELFAPKCSSCKRPVGLGGKGVSFEDRHHNHCFCARCSSTSLVGQGFVPDGDQVLCQ 273
Db 670 NLFSTKCHGCDPVE--AGDKFIEALGHTWHTDTCFICAVCHVNLEQGFYSSKDRPLCK 726

RESULT 8
US-10-374-979-102
; Sequence 102, Application US/10374979
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
```



; Sequence 913, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 913

; LENGTH: 615

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-913

Query Match 16.2%; Score 265; DB 6; Length 615;

Best Local Similarity 29.9%; Pred. No. 2.6e-16;

Matches 55; Conservative 31; Mismatches 88; Indels 10; Gaps 6;

QY 101 CSACGETVMPGSRKLEYG---GQTWHEHCFLCSGCEQPLGSRSFVPDKGAHYCVPCYEN 156

DB 186 CRECKQIGGGDIANFASRAGLACWHPQCFTCTCQELLVLIIFYHVGKVCYGRHAE 245

QY 157 KFAPCARCSTLTGGVTYRD-QPWHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGE 215

DB 246 CLRPRCAQCEIIFSPCTEAEGRHWMDFCCFECEASLGORYVMQRSPHCACAYEA 305

QY 216 LEAPKSSCKRPVGLGGKYVSFEDRHH--HNCFSARCSTSLVGQGFVDPGDQVLC- 272

DB 306 RHAECYCDGGEHI-GLDQSQ-MAYEGQHWASDRCCSRGALLGPRFLPRGLIFCS 363

QY 273 QGCS 276

DB 364 RACS 367

RESULT 12

US-10-408-765A-1767

; Sequence 1767, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1767

; LENGTH: 539

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-1767

Query Match 16.0%; Score 262; DB 6; Length 539;

Best Local Similarity 27.2%; Pred. No. 4.2e-16;

Matches 67; Conservative 37; Mismatches 100; Indels 42; Gaps 10;

QY 57 EDRHFHEGFCRCRCQRSIADPFTCCQDSEILCNDYCYSQCSAC-----GEFVMPG 111

DB 57 EDRHFHEGFCRCRCQRSIADPFTCCQDSEILCNDYCYSQCSAC-----GEFVMPG 111

DB 6 ODKYFIHKCFVCKACGCDLAEGGFFVROGEYICTLDYQRYLGRFCFSCDQFIERGEVV--- 62

QY 112 SRKLEYGGOTWHEHCFLCSGCEQPLGSRSFVPDKGAH-YCVPC-----YENKFA 160

DB 63 ----SALGKTYHPDCFCVCAVCRLPFPFGDRVTFNGRECMCKCSLPSVSGSSAHLQGLR 118

QY 161 RCARCSKTLTQG-GVTYRDQPHRECLVCTGCTPLAGQQFTSRDEDPYCVACFGE 219

DB 119 SCGGCGTEIKNGOALVALDKHMLGCFCKKSC-GKLLNAEYISKDGLPYCEADYHAKFGI 177

QY 220 KSSCKRPVGLGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPGDQVLCQ----- 274

DB 178 RCDSCKEYIT---GRVLEAGEKHYPSCALCVRC-----GOMFA-EGEEMYLQGGSIWH 227

QY 275 --CSQA 278

DB 228 PACRQA 233

RESULT 13

US-10-275-595A-21

; Sequence 21, Application US/10275595A

; GENERAL INFORMATION:

; APPLICANT: YUE, Henry

; APPLICANT: TANG, Y. Tom

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: LU, Dyung Aina M.

; APPLICANT: BAUGHN, Mariah R. L.

; APPLICANT: HILLMAN, Jennifer L.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LAL, Preeti

; APPLICANT: YAO, Monique G.

; APPLICANT: BANDMAN, Olga

; APPLICANT: BURFORD, Neil

; APPLICANT: BATRA, Sajeev

; APPLICANT: KEARNEY, Liam

; APPLICANT: POLICKY, Jennifer L.

; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS

; FILE REFERENCE: PF-0772 USN

; CURRENT APPLICATION NUMBER: US/10/275,595A

; PRIOR FILING DATE: 2003-06-13

; PRIOR APPLICATION NUMBER: US 60/201,960

; PRIOR FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: US 60/202,729

; PRIOR FILING DATE: 2000-05-08

; PRIOR APPLICATION NUMBER: US 60/209,705

; PRIOR FILING DATE: 2000-06-05

; PRIOR APPLICATION NUMBER: US 60/210,149

; PRIOR FILING DATE: 2000-06-07

; PRIOR APPLICATION NUMBER: US 60/213,215

; PRIOR FILING DATE: 2000-06-21

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PERL Program

; SEQ ID NO 21

; LENGTH: 212

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: Incyte ID No: 2913976CDI

US-10-275-595A-21

Query Match 15.9%; Score 260; DB 6; Length 212;

Best Local Similarity 29.5%; Pred. No. 2.4e-16;

Matches 56; Conservative 26; Mismatches 96; Indels 12; Gaps 5;

QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFLCSGCEQPLGSRSFVPDKGAHYCVPCYENKFA 160

DB 22 CAFCHTVPFRELAVEAMKRQYHAQCFTCTCRRLAGQSFYKQDGRPLCEPCYQDTL-E 80

QY 161 RCARCSKTLTQGGVTYRDQPHRECLVCTGCTPLAGQQFT-SRDEDPYCVACFGE 219

DB 81 RCKGCVVRDHIIRALGOAFHPSCFTCVTCARCIGDESPALGSONEVYCLDDFYRKFA 140



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:54:46 ; Search time 40 Seconds  
(without alignments)  
673.181 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCAKNESLYGRKYI.....QGFPDGDQVLCQGCQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1569	96.0	280	T09504	LIM-protein FHL3,
2	968	59.2	279	JC6565	four-and-a-half LI
3	922	56.4	525	T21357	hypothetical prote
4	844	51.6	284	JC7686	activator of CAMP-
5	755.5	46.2	280	G01884	LIM protein FHL-1,
6	730	44.6	454	T21358	hypothetical prote
7	728.5	44.6	280	G02741	skeletal muscle LI
8	718	43.9	279	JG0164	LIM protein, FHL4
9	311.5	19.1	444	A55071	hydrogen peroxide-
10	306.5	18.7	314	JC2324	LIM protein - huma
11	304.5	18.6	557	A55933	paxillin - human
12	303.5	18.6	559	B55933	paxillin - chicken
13	300.5	18.4	256	E88469	protein C2H8.6 [i
14	297.5	18.2	348	T16076	LIM protein - Caen
15	293	17.9	455	A55050	enigma - human
16	286	17.5	1333	T19157	probable metal bin
17	286	17.5	1424	T19156	probable metal bin
18	281.5	17.2	575	C88346	protein F42G4.3a [
19	281.5	17.2	603	T22111	hypothetical prote
20	277.5	17.0	465	T29257	hypothetical prote
21	265.5	16.2	176	T22110	hypothetical prote
22	259	15.8	200	T42678	hypothetical prote
23	257.5	15.7	572	G02845	zyxin - human
24	254	15.5	542	A44358	zyxin - chicken
25	241	14.7	438	T40509	lim domain protein
26	237	14.5	329	T20546	hypothetical prote
27	226	13.8	423	T14842	testin - mouse
28	212	13.0	192	T26544	hypothetical prote
29	195	11.9	194	S41761	cysteine-rich prot

30	195	11.9	194	2	S52335	beta-cysteine-rich
31	195	11.9	194	2	S57472	murine muscle LIM
32	194	11.9	194	2	A55099	muscle LIM protein
33	192.5	11.8	625	2	T20634	hypothetical prote
34	192.5	11.8	639	2	T43190	probable actin-bin
35	185.5	11.3	312	2	T46255	hypothetical prote
36	182	11.1	193	2	S53580	cysteine-rich prot
37	182	11.1	615	2	S06546	finger protein (cl
38	179.5	11.0	3075	2	S14458	laminin alpha-1 ch
39	179	10.9	136	2	T08787	probable testin DK
40	178	10.9	193	2	S12658	cysteine-rich prot
41	177	10.8	192	2	S38879	LIM-domain protein
42	177	10.8	192	2	A49648	cysteine-rich prot
43	174.5	10.7	642	1	JP0079	LIM protein kinase
44	170.5	10.4	367	2	S06582	finger protein (cl
45	169.5	10.4	1017	2	S67804	LRL1 protein - yea

ALIGNMENTS

RESULT 1

T09504

LIM-protein FHL3, skeletal muscle - human

C;Species: Homo sapiens (man)

C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Nov-2000

C;Accession: T09504; PD0036

R;Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 255, 245-250, 1999

A;Title: The LIM proteins FHL1 and FHL3 are expressed differently in skeletal muscle.

A;Reference number: PD0036; MUID:99160848; PMID:10049693

A;Accession: T09504

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-280 <MOR>

A;Cross-references: EMBL:U60116; NID:g4116529; PIDN:AAC04466.2; PID:g4416530

A;Accession: PD0036

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-39 <MO2>

A;Cross-references: GB:U60116

C;Genetics:

A;Map position: 1

C;Keywords: metal binding; zinc finger

Query Match 96.0%; Score 1569; DB 2; Length 280;  
Best Local Similarity 95.7%; Pred. No. 6.9e-106;  
Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy	1	MSESFDCAKNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFYEDRH	60
Db	1	MSESFDCAKNESLYGRKYIQTDSGPCYCPYDNTFANTCAECQQLIGHDSRELFYEDRH	60
Qy	61	FHEGFCRCRCQRSLADEPFTCCQSELNDCYCSAFSSQCSAGETVMPGSRKLEYGGQ	120
Db	61	FHEGFCRCRCQRSLADEPFTCCQSELNDCYCSAFSSQCSAGETVMPGSRKLEYGGQ	120
Qy	121	TWHEHCFCLSCGCEPLGSRFVPDKGAHYCPVCYENKFAPCARCSKTLTGGGVYRDQP	180
Db	121	TWHEHCFCLSCGCEPLGSRFVPDKGAHYCPVCYENKFAPCARCSKTLTGGGVYRDQP	180
Qy	181	WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFELFAPKCSSCKRPTVLGGGKYVSFE	240
Db	181	WHRECLVCTGCTPLAGQQFTSRDEDPYCVACFELFAPKCSSCKRPTVLGGGKYVSFE	240
Qy	241	DRWHHNCFCARCSSTLVGGQFVDPDGDQVLCQGCQAGP	280
Db	241	DRWHHNCFCARCSSTLVGGQFVDPDGDQVLCQGCQAGP	280

RESULT 2

JC6565

four-and-a-half LIM-only protein 2 - human

C:Species: Homo sapiens (man)  
C>Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jul-2000  
C:Accession: JC6565  
R:Chan, K.K.; Tsui, S.K.W.; Lee, S.M.Y.; Luk, S.C.W.; Liew, C.C.; Fung, K.P.; Waye, M.M.  
Gene 210, 345-350, 1998  
A:Title: Molecular cloning and characterization of FHL2, a novel LIM domain protein pref  
A:Reference number: JC6565; MUID:98248917; PMID:9573400  
A:Accession: JC6565  
A:Molecule type: mRNA  
A:Residues: 1-279 <CHA>  
A:Cross-references: GB:U29332; NID:gl845201; PIDN:AAC52073.1; PID:gl377897  
A:Experimental source: heart  
C:Genetics:  
A:Map position: 2q12-2q13  
C:Superfamily: LIM metal-binding repeat homology  
C:Keywords: cardiac muscle; heart; zinc finger  
F:221-275/Domain: LIM metal-binding repeat homology <LIM>

Query Match 59.2%; Score 968; DB 2; Length 279;  
Best Local Similarity 53.4%; Pred. No. 1e-62;  
Matches 148; Conservative 59; Mismatches 70; Indels 0; Gaps 0;

QY 1 MSEPFDCAKNSLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFEYDRH 60  
DB 1 MTERFDCHNCNSLFGKYLILRESPYCVVCFETLFTANTCEGCRPIGDCDKDLTKDRH 60

QY 61 FHEGFCRCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGQ 120  
DB 61 WHEACFHCSCQNSLVDKFAAKEDEQLLDCYSNEYSSKQCECKTTPGTRKMEYKGS 120

QY 121 TWHEHCFCLSGCEQPLGSRFVDPDKGAHYCPYENKFAAPRCARCKTLTQGGVYRDOP 180  
DB 121 SWHETCFICHCQQPPIGTGKSFIPKDNQNFVCPYKQHAMQCVQCKMPITTTGGVYREOP 180

QY 181 WHRECLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVSPE 240  
DB 181 WKECFVCTACKQQLSGQRTARDDFAYCLNCFGLDIAKKACAGCTNPISGLGGTKYISPE 240

QY 241 DRHWHNCFSCARCSTSLVGQGFVDPDQVLCQCGSQ 277  
DB 241 ERQWHDNCFNCKKSLSLVGRGFLTERDILCPDCGK 277

RESULT 3  
T21357  
hypothetical protein F25H5.1a - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21357  
R:Steward, C.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19412  
A:Accession: T21357  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-525 <WIL>  
A:Cross-references: EMBL:281068; PIDN:CAB02980.1; GSPDB:GN00019; CESP:F25H5.1a  
A:Experimental source: clone F25H5  
C:Genetics:  
A:Gene: CESP:F25H5.1a  
A:Map position: 1  
A:Introns: 35/3; 114/3; 153/3; 186/3; 253/3; 288/1; 312/1; 431/1; 478/3

Query Match 56.4%; Score 922; DB 2; Length 525;  
Best Local Similarity 50.4%; Pred. No. 3.4e-59;  
Matches 139; Conservative 64; Mismatches 73; Indels 0; Gaps 0;

QY 2 SESEFDCAKNSLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFEYDRH 61  
DB 203 SDHFCWCQDQTLTGQYIMRDEQPYCIKCYEDVFANQCDCAKPIGIDSKDLTKYKHW 262

QY 62 HEGCFRCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGQ 121  
DB 62 WHRECLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVSPE 240

DB 263 HEHCFCLSCMCKISLVDMPFGSKNDRIFLFCNSCYDQAFATRCDCGNEIFRAGMKMEYKKGQ 322  
QY 122 WHEHCFCLSGCEQPLGSRFVDPDKGAHYCPYENKFAAPRCARCKTLTQGGVYRDQPW 181  
DB 323 WIDKFCFCAHCKLAIGTKSFIPKNDVDFCGPYEKEFATRCCKKVIITAGGVTYKNEPW 382  
QY 182 HRECLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVSPE 241  
DB 383 HRECFCCITNCNSSLAGORETSKDEPYCANCYGLDFAKRCNACTPITIGGAKFISPED 442

QY 242 RIWHNCFSCARCSTSLVGQGFVDPDQVLCQCGSQ 277  
DB 443 RHWHNDFCAOCTTSLVGKGFITDGHGHEILCPECAK 478

RESULT 4  
JC7686  
activator of cAMP-responsive element modulator, testis - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: JC7686  
R:Palermo, I.; Litrico, L.; Emmanuele, G.; Giuffrida, V.; Sassone-Corsi, P.; De Cesar  
Biochem. Biophys. Res. Commun. 283, 406-411, 2001  
A:Title: Cloning and expression of activator of CREM in testis in human testicular ti  
A:Reference number: JC7686; MUID:21226154; PMID:11327716  
A:Contents: Testis  
A:Accession: JC7686  
A:Molecule type: mRNA  
A:Residues: 1-284 <PAL>  
A:Cross-references: GB:AF278541  
C:Comment: This protein activates cAMP-responsive element modulator in a phosphorylat  
of human spermatogenesis.  
C:Genetics:  
A:Gene: act  
A:Map position: 6q  
A:Introns: 53/3; 112/3; 168/3; 231/3  
C:Keywords: spermatogenesis; testis

Query Match 51.6%; Score 844; DB 2; Length 284;  
Best Local Similarity 47.8%; Pred. No. 8.6e-54;  
Matches 130; Conservative 51; Mismatches 91; Indels 0; Gaps 0;

QY 5 FDCAKNSLYGRKYIQTDSGPYCPYDNTFANTCAEQQLIGHDSRELFEYDRH 64  
DB 6 FYCQYCTASLLGKKYLVKVDSPYCVTYDRVFSNYECKKPTESDKDFCYKRVHHEG 65

QY 65 CFRCCRCQRLADEPFTQDSSELLCNDYCYSFSSQCSACGETVMPGSRKLEYGGOTWHE 124  
DB 66 CFKCTKCNHSLVEKFPAAKDERLLCTECYSNCCSSKCFHCKRTIMPGRMMEFKGNWHE 125

QY 125 HCFLSGCEQPLGSRFVDPDKGAHYCPYENKFAAPRCARCKTLTQGGVYRDQPW 184  
DB 126 TCFVCNCRQPIGTGKPLISKESGNCVCPCFEKEFAHCNFKKVTSGTTCDDLWHE 185

QY 185 CLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPIVGLGGKVSPE 244  
DB 186 CFLSGCRKDLCEEQFMSRDDYFPCMDYCNHLYANKVACSKPISGLTCAKFCFQDSQW 245

QY 245 HHNCFSCARCSTSLVGQGFVDPDQVLCQCGSQ 276  
DB 246 HSECFNCRGCSVSLVGKGFITDGHGHEILCPECAK 478

RESULT 5  
G01884  
LIM protein FHL-1, cardiac - human  
C:Species: Homo sapiens (man)  
C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 16-Jul-1999  
C:Accession: G01884  
R:Waye, M.M.Y.  
submitted to the EMBL Data Library, June 1995  
A:Reference number: H00697  
A:Accession: G01884

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-280 <SWA>  
 A:Cross-references: EMBL:U29538; NID:gl1377898; PID:gl1377899  
 C:Genetics:  
 A:Gene: FHL-1

C:Superfamily: LIM metal-binding repeat homology  
 C:Keywords: cardiac muscle; duplication; heart  
 F:40-92/Domain: LIM metal-binding repeat homology <LIM1>  
 F:101-153/Domain: LIM metal-binding repeat homology <LIM2>  
 F:162-212/Domain: LIM metal-binding repeat homology <LIM3>  
 F:221-276/Domain: LIM metal-binding repeat homology <LIM4>

Query Match 46.2%; Score 755.5; DB 2; Length 280;  
 Best Local Similarity 44.2%; Pred. No. 1.9e-47;  
 Matches 123; Conservative 50; Mismatches 104; Indels 1; Gaps 1;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFLYEDRH 60  
 DB 1 MAEKFDCHYCRVPLQGGKYVQKHCCCLKCFKANTVCBCRPFGADSKVEHYKNRF 60  
 QY 61 FHEGFCRCRCORSLEADPEFTQDSLELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 DB 61 WHDTCFERCAKCLHPLANETFAVDKILCNKCTTREDSSCKGCFKAI VAGDQNVYKGT 120  
 QY 121 TWHEHCFCLSGCGEPLGSRSFVPDKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRDQP 180  
 DB 121 VWHKDCFTCSNCKQVIGTSGFPFGEDFYCVTCHETKFAKHCVKCNKAITSGGIYQDQP 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLG-GGKYVSF 239  
 DB 181 WHADCFVVCVTCSSKLLAGAAFTVEDQYCYVDFKFAKCKAGCNKPIITGFGKGSVVAY 240  
 QY 240 EDRHWHNCFSCARSTSLVGQGFVDPDGVLCQCGSQ 277  
 DB 241 EGQSHDYCFHCKKCSVNLANKRFVHQEQVYCPDCAK 278

## RESULT 6

T21358  
 hypothetical protein F25H5.1b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T21358

R:Steward, C.  
 submitted to the EMBL Data Library, October 1996

A:Reference number: Z19412  
 A:Accession: T21358  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-454 <WIL>  
 A:Cross-references: EMBL:281068; PIDN:CAB02981.1; GSPDB:GN00019; CESP:F25H5.1b  
 A:Experimental source: Clone F25H5  
 C:Genetics:  
 A:Gene: CESP:F25H5.1b  
 A:Map position: 1  
 A:Introns: 35/3; 114/3; 153/3; 186/3; 253/3; 288/1; 312/1

Query Match 44.6%; Score 730; DB 2; Length 454;  
 Best Local Similarity 48.5%; Pred. No. 1.9e-45;  
 Matches 111; Conservative 51; Mismatches 67; Indels 0; Gaps 0;

QY 2 SESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFLYEDRH 61  
 DB 203 SHFCWCQDQFTLGRIYMRDEQPYCIKYEDVFANQCDCAKPIGIDSKDLSDKDRHW 262  
 QY 62 HEHCFRCRCORSLEADPEFTQDSLELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGGT 121  
 DB 263 HEHCFCLSMCKISLVDMPSGKNDRIFCNSCYDQAFATRCDCGNEIFRAGMKMEYKQKQ 322  
 QY 122 WHRECLVCTGCGEPLGSRSFVPDKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRDQP 181  
 DB 323 WHDKFCFCAHCKLAIGTKSFIPKNDVFCGPGCYEKEKFAFRCCKKVIITAGGVYTKNEPW 382

QY 182 HRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIG 230  
 DB 383 HRECFCTNCNSSLAGQRTSKDEKPYCANCYGDLEFAKRCNACTKPIIG 431

## RESULT 7

G02741

skeletal muscle LIM-protein SLIM1 - human

C:Species: Homo sapiens (man)  
 C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 16-Jul-1999  
 C:Accession: J04893; G02741

R:Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 225, 632-638, 1996

A:Title: SLIM defines a novel family of LIM-proteins expressed in skeletal muscle.

A:Reference number: J04893; MUID:96354835; PMID:8753811

A:Accession: J04893

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-280 <MO2>

A:Cross-references: EMBL:U60115; NID:gl1381807; PID:gl1381808

C:Superfamily: LIM metal-binding repeat homology

F:162-212/Domain: LIM metal-binding repeat homology <LIM>

Query Match 44.6%; Score 728.5; DB 2; Length 280;  
 Best Local Similarity 42.8%; Pred. No. 1.7e-45;  
 Matches 119; Conservative 50; Mismatches 108; Indels 1; Gaps 1;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFLYEDRH 60  
 DB 1 MAEKFDCHYCRVPLQGGKYVQKHCCCLKCFKANTVCBCRPFGADSKVEHYKNRF 60  
 QY 61 FHEGFCRCRCORSLEADPEFTQDSLELLCNDVCYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 DB 61 WHDTCFERCAKCLHPLANETFAVDKILCNKCTTREDSSCKGCFKAI VAGDQNVYKGT 120  
 QY 121 TWHEHCFCLSGCGEPLGSRSFVPDKGAHYCPVCYENKFAPRCARCSKTLTQGGVYRDQP 180  
 DB 121 VWHKDCFTCSNCKQVIGTSGFPFGEDFYCVTCHETKFAKHCVKCNKAITSGGIYQDQP 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLG-GGKYVSF 239  
 DB 181 WHADCFVVCVTCSSKLLAGAAFTVEDQYCYVDFKFAKCKAGCNKPIITGFGKGSVVAY 240

## RESULT 8

JG0164

LIM protein, FHL4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 11-May-2000

C:Accession: JG0164

R:Morgan, M.J.; Madgwick, A.J.A.

Biochem. Biophys. Res. Commun. 255, 251-255, 1999

A:Title: The fourth member of the FHL family of LIM proteins is expressed exclusively

A:Reference number: JG0164; MUID:99160849; PMID:10049694

A:Accession: JG0164

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-279 <MOR>

A:Cross-references: GB:AF053486

Query Match 43.9%; Score 718; DB 2; Length 279;  
 Best Local Similarity 42.4%; Pred. No. 9.6e-45;  
 Matches 118; Conservative 46; Mismatches 112; Indels 2; Gaps 2;

QY 1 MSESFDCAKCNESLYGRKVIQTDSPGYCPVPCYDNTFANTCAEQQLIGHDSRELFLYEDRH 60  
 DB 1 MSE-FKCHCEESLOGKKYVQKHCCCLKCFKANTVCBCRPFGADSKVEHYKNRF 59

```
Qy 61 PHECFRCRCQRLADEPFTQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGGQ 120
Db 60 WHNTFCQTKSQLLATETVAMDKNILCNKCATRVTFPKCXGCKLXIEGDHXYEYKGS 119
Qy 121 TWHEHCFCLSCGEQPLGSRFVDPKGAHYCVPCYENKFAFRCARCSKTLTQGGVYRDQP 180
Db 120 IWHKNCVFCNTCKDIITGKNFFPKDEGFYCVCTYDALFTKCMCKCKPITSGGVSYQDQP 179
Qy 181 WHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPKCSSCKRPIVGLG-GGKYVSF 239
Db 180 WHSECFVVCSSKSLSQRTAMDDQYFCVDYKYNIAKCKAGCKNITGFGKGANVVAH 239
Qy 240 EDRHWHHNCFCARCSTSLVGQGFVDPDQVLCQCSQ 277
Db 240 EQNSWHDYCFNCKTCSVNLANKHEVFHDEQVYCPDCAR 277

RESULT 9
A55071
hydrogen peroxide-inducible protein hic-5 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A55071
R:Shibanuma, M.; Mashimo, J.; Kuroki, T.; Nose, K.
J. Biol. Chem. 269, 26767-26774, 1994
A:Title: Characterization of the TGFbeta1-inducible hic-5 gene that encodes a putative
A:Reference number: A55071; MUID:95014536; PMID:7929412
A:Accession: A55071
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-444 <SHI>
C:Superfamily: LIM metal-binding repeat homology
F:211-261/Domain: LIM metal-binding repeat homology <LIM1>
F:270-320/Domain: LIM metal-binding repeat homology <LIM2>
F:329-379/Domain: LIM metal-binding repeat homology <LIM3>
F:388-438/Domain: LIM metal-binding repeat homology <LIM4>

Query Match 19.1%; Score 311.5; DB 2; Length 444;
Best Local Similarity 32.4%; Pred. No. 2.2e-15;
Matches 61; Conservative 23; Mismatches 89; Indels 15; Gaps 3;

Qy 101 CSAGETVMPGSRKLEYGGQTWHEHCFCLSCGEQPLGSRFVDPKGAHYCVPCYENKFAF 160
Db 211 CGSCKNKEI--AGQVVTALGRAWPEHFLHSCGSCSTTLGSSFFEKDGAPCECYFERFSP 268
Qy 161 RCARCSKTLTQGGVYRDQPHHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPK 220
Db 269 RCGFCNQPIRHKMVTALGTHWHPHFHCCVSCGEPFGEFGEFHERGRPYCRRDFLQLFAPR 328
Qy 221 CSSCKRPIVGLGGKYVSFEDRHHHNCFCARCSTSLVGQGFVDPDQVLCQG----- 274
Db 329 CQCGQGPIL----DNYISALSWHPDPCVFCRECLAPFGSGSFHEGRPLCENHFAQR 384
Qy 275 ---CSQAG 279
Db 385 GSLCATCG 392

RESULT 10
JC2324
LIM protein - human
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: JC2324
R:Rearden, A.
Biochem. Biophys. Res. Commun. 201, 1124-1131, 1994
A:Title: A new LIM protein containing an autoeoptope homologous to "senescent cell antigen
A:Reference number: JC2324; MUID:94296379; PMID:7517666
A:Accession: JC2324
A:Molecule type: mRNA
A:Residues: 1-314 <REA>
A:Crogs-references: EMBL:U09284; NID:g516011; PIDN:AAA20086.1; PID:g516012
```

```
A:Experimental source: liver
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: glycoprotein
F:10-62/Domain: LIM metal-binding repeat homology <LIM1>
F:71-121/Domain: LIM metal-binding repeat homology <LIM2>
F:135-184/Domain: LIM metal-binding repeat homology <LIM3>
F:149-157/Region: PINCH autoeoptope
F:193-243/Domain: LIM metal-binding repeat homology <LIM4>
F:252-303/Domain: LIM metal-binding repeat homology <LIM5>
F:87/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.7%; Score 306.5; DB 2; Length 314;
Best Local Similarity 27.4%; Pred. No. 3.9e-15;
Matches 77; Conservative 40; Mismatches 145; Indels 19; Gaps 7;

Qy 3 ESFDCAKCNESLYGRKYIQTDSGPYCVPCYDNTFANTCAEQQLIGHDSRLFLYEDRHIF 62
Db 34 QCFVCAQCFQFPGLFYEFGRKYCEHDFQMLFAPCCHQCGEFI--IGRVIKANNWSH 91
Qy 63 EGFCRCRCQRLADEPFTQDSSELLCNDYCSAFSSQ-----CSAGETVMPGSRKLEY 117
Db 92 PECFRCDCQEVLDIGFVKNAGRHLCRPNCHNRKARGLKYICQKCHAIL--DEQPLIF 149
Qy 118 GQGTWHEHCFCLSCGEQPL--GSRFVDPKGAHYCVPCYENKFAFRCARCSKTLTQGGV 175
Db 150 KNDPYHPDHFNCAKCGKELTADAREL--KGELYCLPCHDKMGVPICGACRRPIEGRVYN 206
Qy 176 YRQDPWHRECLVTCGOTPLAQOQTSRDEDPYCVACFGELEFAPKCSSCKRPIVGLGGK 235
Db 207 AMGQKQWHEHFVCAKCEKPFGLHRRHYERKGLAYETHYNQLFGDVCFHCNRVI----EGD 262
Qy 236 YVSEFEDRHHHNCFCARCSTSL-VGQGFVDPDQVLCQCG 275
Db 263 VVSALNKAQVNCVFCACSTCNKTLTKNKFVEFDMKPKVCKKC 303

RESULT 11
A55933
paxillin - human
N:Alternate names: peroxisomal membrane protein 3
C:Species: Homo sapiens (man)
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 20-Apr-2000
C:Accession: A55933
R:Salgia, R.; Li, J.L.; Lo, S.H.; Brunkhorst, B.; Kansas, G.S.; Sobhany, E.S.; Sun, Y.
J. Biol. Chem. 270, 5039-5047, 1995
A:Title: Molecular cloning of human paxillin, a focal adhesion protein phosphorylated
A:Reference number: A55933; MUID:95197488; PMID:7534286
A:Molecule type: mRNA
A:Residues: 1-557 <SAL>
A:Crogs-references: GB:U14588; NID:g704347; PIDN:AAC50104.1; PID:g704348
C:Genetics:
A:Gene: GDB:PXN
A:Crogs-references: GDB:702105
A:Map position: 12q24-12q24
C:Superfamily: LIM metal-binding repeat homology
C:Keywords: cell adhesion; cytoskeleton; peroxisome; phosphoprotein; zinc finger
F:46-55/Region: proline-rich
F:324-374/Domain: LIM metal-binding repeat homology <LIM1>
F:383-433/Domain: LIM metal-binding repeat homology <LIM2>
F:442-492/Domain: LIM metal-binding repeat homology <LIM3>
F:501-551/Domain: LIM metal-binding repeat homology <LIM4>
F:31,118,181/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 18.8%; Score 304.5; DB 2; Length 557;
Best Local Similarity 27.5%; Pred. No. 8.4e-15;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

Qy 40 CAEQQLIGHDSRLFLYEDRHIFHGCGFCRCRCQRLADEPFTQDSSELLCNDYCSAFSS 99
Db 324 CGACKKPFA--GQVVTAMGTWHPHFVCTHCQCEIGSRNFFERDGGPYCEKDYHNLFSP 381
Qy 100 QCSAGETVMPGSRKLEYGGQTWHEHCFCLSCGEQPLGSRFVDPKGAHYCVPCYENKFA 159
```



```

Db      382 RCYCYNGNPL--DKVVTALDRTWHPHFCAQCGAFFPGGFHEKDGKAYCRKDYFDMAF 439
QY      160 PRCARCSKTLTGGVYTRDQDPWHRECLVCTGCQTPLAQOQFTSRDEDPYCVACFGELFAP 219
Db      440 PKCGGCARAILENYISALNTLWHPECFVCRCECTPFYFNGSFFEHDPGQYCEVHYHERRGS 499
QY      220 KCSSCKRPVIGLGGKYVYFDRWHHNCFSARCSTSLVGQGFVPDQGVLCQGC 275
Db      500 LCSGCKPKPT----GRCITAMAKKFPHPHFCAFLKQLNKGTFKEQNDKPYCQNC 551

RESULT 12
paxillin - chicken.
C;Species: Gallus gallus (chicken)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C;Accession: B55933
R;Saigita, R.; Li, J. L.; Lo, S. H.; Brunkhorst, B.; Kansas, G. S.; Sobhany, E. S.;
J. Biol. Chem. 270, 5039-5047, 1995
A;Title: Molecular cloning of human paxillin, a focal adhesion protein phospho-
A;Reference number: A55933; MUID:95197488; PMID:7534286
A;Accession: B55933
A;Molecule type: mRNA
A;Residues: 1-559 <SAL>
A;Cross-references: GB:U14589; NID:g704349; PIDN:AAC59665.1; PID:g704350
C;Superfamily: LIM metal-binding repeat homology
C;Keywords: cell adhesion; cytoskeleton; phosphoprotein; zinc finger
F;46-55/Region: proline-rich
F;326-376/Domain: LIM metal-binding repeat homology <LIM1>
F;385-435/Domain: LIM metal-binding repeat homology <LIM2>
F;444-494/Domain: LIM metal-binding repeat homology <LIM3>
F;503-553/Domain: LIM metal-binding repeat homology <LIM4>
F;31,118/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match      18.6%; Score 303.5; DB 2; Length 559;
Best Local Similarity 27.5%; Pred. No. 9.9e-15;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3

QY      40 CAEQCOLIGHDSRELFEYDRHFHEGCFCCRCORSLADEPFTQDSLELLNCDCYCSAFSS 99
Db      326 CGACKRKPTA--GQVVTAMGKTWHPHFVCTHCQEEIGSRNFFERDGPYCEKDYHNLFSF 383
QY      100 QCSAGCETVMPGSRKLEYGGQVTWHEHCFCLSCGCEQPLGSRSFVPDKGAHVCPVCYNKFA 159
Db      384 RCYCYNGNPL--DKVVTALDRTWHPHFCAQCGVFFGPGGFHEKDGKAYCRKDYFDMAF 441
QY      160 PRCARCSKTLTGGVYTRDQDPWHRECLVCTGCQTPLAQOQFTSRDEDPYCVACFGELFAP 219
Db      442 PKCGGCARAILENYISALNTLWHPECFVCRCECTPFYFNGSFFEHDPGQYCEVHYHERRGS 501
QY      220 KCSSCKRPVIGLGGKYVYFDRWHHNCFSARCSTSLVGQGFVPDQGVLCQGC 275
Db      502 LCSGCKPKPT----GRCITAMAKKFPHPHFCAFLKQLNKGTFKEQNDKPYCQNC 553

```

RESULT 13  
E88469  
protein C28H8.6 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: E88469  
R/anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigation  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999  
A/Accession: E88469  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-256 <STO>  
A/Cross-references: GB|chr\_III; PIDN:AAA62294.1; PID:9669012; GSPDB:GN00021; C

```

Db      382 RCYYCNGPIL--DKVVTALDRTHWPEHFCQAQCGAFTFGPEGFHEKDGKAYCRKDYFDMEA 439
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY      160 PRCARGSKTLTQGGVTVYRDQWPHRECLVCTGCQTPLAGQQFTSRDEDPCYVACFGELFAP 219
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      440 PKCGGCARAILENVISALNTLWPECFVCRECFPFVNGSFFFDHQGPYCEVHYHERRGS 499
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY      220 KCSCKRPVGLGGGKYVSPEDRRHHNHCFSARCSTSLVGQGVDPDQVLCQGC 275
      || :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db      500 LCSGCKRPIT---GRCLTAMAKKFFHEFVCAFLKQLNKGTFKQNDKPYCQNC 551
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

RESULT 12
B55933
paxillin - chicken.
C/Species: Gallus gallus (Chicken)
C/Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 21-Jul-2000
C/Accession: B55933
R/Salgia, R.; Li, J.L.; Lo, S.H.; Brunkhorst, B.; Kansas, G.S.; Sobhany, E.S.; Sun, Y.;
J. Biol. Chem. 270, 5039-5047, 1995
A/Title: Molecular cloning of human paxillin, a focal adhesion protein phosphorylated by
A/Reference number: A55933; PMID:95197488; PMID:7534286

```

```

A: Cross-references: GB:U14589; NID:g704349; PIDN:AC59665.1; PID:g704350
C: Superfamily: LIM metal-binding repeat homology
C: Keywords: cell adhesion; cytoskeleton; phosphoprotein; zinc finger
F: 46-55/Region: proline-rich
F: 326-376/Domain: LIM metal-binding repeat homology <LIM1>
F: 385-435/Domain: LIM metal-binding repeat homology <LIM2>
F: 444-494/Domain: LIM metal-binding repeat homology <LIM3>
F: 503-553/Domain: LIM metal-binding repeat homology <LIM4>
F: 31,118/Binding site: phosphate (Tyr) (covalent) $status predicted

Query Match          38.6%; Score 303.5; DB 2; Length 559;
Best Local Similarity 27.5%; Pred. No. 9.9e-15;
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3

Qy  40  CAECQQLIGHDSRELFLYEDRIHFHGGCFKCCRCQRSLADEPFTQDSSELLNCDCYCSAFSS 99
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   326  CGACKKPIA--GQVVTAMGKTWHPEHFVCTHCQEIGSRNFFERDGPQYCEKDYHNLFP 383
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   100  QCSAGETVMPGSRKLEYGGOTWIEHCFCLSCGCEQPLGSRSFVDPDKGAHVCPVCYENKFA 159
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   384  RYVYCNGLP--DKVVTALDRTWIPEHFFCAQGVFFGPGFHEKGKAYCRKDYDFDMA 441
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   160  PRCAKCSKTLTGGVTVYRDQPPWHRECLVCTGCQPTLAGOQFTSRDEPDYCVACFGELFAP 219
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   442  PKCGCARAILENYISALNTLWHPECFVRCRECTPFINGSFPEHDGPQYCEVHYHERRGS 501
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   220  KCSSCKRPVILGGGKYVYFEDRIWHHNNCFSCARCSTSLVGQGVFVDPDGOVLQCGC 275
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   502  LCGGCKQKIT---GRCTIAMKGFHPHGFYCAFLQKLNKGTFKKQNDKPYCQC 553
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

RESULT 13  
E88469  
protein C28H8.6 [imported] - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: E88469  
R;anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes  
A/Reference number: A75000; PMID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2104, 1999.  
A/Accession: E88469  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-256 <STO>  
A/Cross-references: GB:chr\_III; PIDN:AAA62294.1; PID:g669012; GSPDB:GNO00021; CESP:C28H8.6

```
C:Genetics:  
A:Gene: C28H8.6  
A:Map position: 3
```

Query Match            18.4% ; Score 300.5 ; DB 2 ; Length 256;  
Best Local Similarity 30.6% ; Pred. No. 8.9e-15;  
Matches 57; Conservative 27; Mismatches 87; Indels 15; Gaps 3;

Qy	101	CSACGETVMGSRKLEYGGQTWIEHCFLCSCGCQPGLGRSFVPDKGAHYCVPCYENKFAP	160
Db	19	CAACGKPDI--GVVIALAGKMWPHEHYTCCECGAEJGQRFFERNRGAFCEDYHNQFSP	76
Qy	161	RCARCSXTLQGQGYTYRDQPPWHRECLVTCTGCPQLAGQQFTSDEDPYCVCAGELFAPK	220
Db	77	KCOGCHRAITDRCVSVMNKNFHIECTFCAECNQPFEGDGFEHKNGQTYCKRDFRLFAPK	136
Qy	221	CSSCKRPVLGGKKYYVFSDRRHHNHCFSCARCSTSLVGQGFVPDGDVLCOG-----	274
Db	137	CNGSQBIT----SNFITALGTWHHPDCFYCQHCGVSVFNCGASFEHNGCAPLCCRHVHESR	192
Qy	275	---CSQ 277	
Db	193	GSI CSQ 198	

RESULT 14  
 LIM676  
 LIM protein - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: T16076; T42393  
 R:Minx, P.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid F14D12.  
 A:Reference number: z18457  
 A:Accession: T16076  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-348 <MIN>  
 A:Cross-references: EMBL:U41021; NID:g10866679; PID:g1086681; PIDN:AA82334.1; C:  
 R:Hober, O.; Moerman, D.G.; Clark, K.A.; Beckerle, M.C.; Ruvkun, G.  
 J. Cell Biol. 144, 45-57, 1999  
 A:Title: A conserved LIM protein that affects muscular adhesion junction integr  
 A:Reference number: Z22152; MUID:99102422; PMID:9885243  
 A:Accession: T42393  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-348 <HOB>  
 A:Cross-references: EMBL:AF035583; NID:g5661097; PIDN:AAD09435.1; PID:g26661098  
 C:Genetics:  
 A:Gene: unc-97; F14D12.2  
 A:Introns: 48/2; 86/1; 126/2; 177/2; 258/3; 300/2  
 C:Function:

	Query Match	18.28;	Score 297.5;	DB 2;	Length 348;
	Best Local Similarity	27.08;	Pred. No. 1.9e-14;		
	Matches 73;	Conservative 35;	Mismatches 133;	Indels 29;	Gaps 7;
Qy	2	SESFDCAKNESLYGRKYIQITDSPYCVPCYQDNTFAMTCAECQOLIGHDSRELFYDRHF	61		
Db	44	SECFVCAQCFEPFPDGIYFEVEGRKYCEHDFHVLFSPPCKGCKNEFI--VGRVIKAMNASW	101		
Qy	62	HEGFCRCRCORSIADLPFTQDSSELCNDC-----YCSAFSSOCSAGETVMP	110		
Db	102	HPGFCCEIKNQLADVGLFNAGRALRCRENEREKAAGHGRV-----CHKHAMIDD	155		
Qy	111	GSRKLEYGGQTWHEHCFCLSCGEQPL--GSRSFVPDGAHYVPCYENKFNAPRCARCSKT	168		
Db	156	G-QHIKFGDSFPHYHKCKRCNNELTASREV---NGELYCLRCHDTMGIPITGACHRP	211		
Qy	169	LTOGCVTVYRQDPWHEHRECLVCTGCTPLAGQQFTSRDEPDCVACFGLEFAPKSSCKRPI	228		

Db 212 IEERVIAALGKHVHEHVFVCSVCERKPFGLGHRHYERKGLPYCEQHFKLFGNLFCFCGDPC 271  
QY 229 VGLGGKVVSVFEDRHHHNCFCSCARCSL 258  
Db : : : : :  
272 C---GEVFOALQKTWCVKCFSCFCDDKL 297

## RESULT 15

A55050  
enigma - human  
C:Species: Homo sapiens (man)  
C:Date: 18-Nov-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: A55050  
J. Wu, R.Y.; Gill, G.N.  
J. Biol. Chem. 269, 25085-25090, 1994  
A:Title: LIM domain recognition of a tyrosine-containing tight turn.  
A:Reference number: A55050; MUID:95014287; PMID:7929196  
A:Accession: A55050  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-455 <WUA>  
A:Cross-references: GB:L35240; NID:g561636; PIDN:AAC37565.1; PID:g561637  
C:Superfamily: LIM metal-binding repeat homology; GLGF domain homology  
F:9-81/Domain: GLGF domain homology <GLG>  
F:280-330/Domain: LIM metal-binding repeat homology <LIM1>  
F:339-389/Domain: LIM metal-binding repeat homology <LIM2>  
F:398-450/Domain: LIM metal-binding repeat homology <LIM3>

Query Match 17.9%; Score 293; DB 2; Length 455;  
Best Local Similarity 32.4%; Pred. No. 4.8e-14;  
Matches 56; Conservative 27; Mismatches 86; Indels 4; Gaps 2;  
QY 101 CSACGETVMPGSRKLEYGGQTWHEHCFLCSCGEQPLGSRSFVPDKGAHYCVPCYENKFPAP 160  
Db : : : : :  
280 CHQCHKVIR--GRYLVALGHAYHPEEFVCSQCGKYLEGGFFPEKGAIFCPCYDVRYP 337  
QY 161 RCARCSKTLTQGGVYTRDQPHRECLVCTGCTPLAGQOFTSRDEDPYCVCAGFELFAPK 220  
Db : : : : :  
338 SCACCKKITGTIMHAKMTWHVHCFTCAACKTPIRNRAFYMEGVPCYCDYERKMFGTK 397  
QY 221 CSSCKRPVGLGGKVVSVFEDRHHHNCFCSCARCSLTVGOGFVDPGDQVLQ 273  
Db : : : : :  
398 CHGCDFKID--AGDRFLEALGFSWHDTCFVCAICQINLEGKTFYSKKDRPLCK 448

Search completed: August 22, 2003, 20:00:42  
Job time : 43 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:44:56 ; Search time 23 Seconds  
(without alignments)  
572.499 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCARKNESLYGRKYI.....QGFVPDQVLCQCSQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1578.5	96.5	289	1 SLI2_MOUSE	Q9R059 mus musculus
2	1569	96.0	280	1 SLI2_HUMAN	Q13643 homo sapien
3	998	61.0	279	1 SLI3_RAT	Q35115 rattus norv
4	991	60.6	279	1 SLI3_MOUSE	O70433 mus musculus
5	968	59.2	279	1 SLI3_HUMAN	Q14192 homo sapien
6	762.5	46.6	280	1 SLI1_HUMAN	Q13642 mus musculus
7	760.5	46.5	280	1 SLI1_MOUSE	P37447 mus musculus
8	313	19.1	386	1 LPAX_HUMAN	O60711 homo sapien
9	309.5	18.9	325	1 PINC_HUMAN	P48059 homo sapien
10	305	18.7	495	1 MLP2_DROME	Q24400 drosophila
11	304.5	18.6	591	1 PAXI_HUMAN	P49023 homo sapien
12	303.5	18.6	559	1 PAXI_CHICK	P49024 gallus gall
13	300.5	18.4	256	1 YP96_CAEEL	O09476 caenorhabdi
14	297.5	18.2	348	1 UN97_CAEEL	P50464 caenorhabdi
15	265	16.2	615	1 LM06_HUMAN	Q43900 homo sapien
16	257.5	15.7	572	1 ZYX_HUMAN	Q15942 homo sapien
17	254	15.5	542	1 ZYX_CHICK	O04584 gallus gall
18	244.5	15.0	564	1 ZYX_MOUSE	O62523 mus musculus
19	240	14.7	476	1 TRI6_HUMAN	Q15654 homo sapien
20	237	14.5	329	1 PIN2_CAEEL	P19157 caenorhabdi
21	226	13.8	423	1 TES_MOUSE	P47226 mus musculus
22	220	13.5	421	1 TES_HUMAN	Q90418 homo sapien
23	200	12.2	194	1 CSR3_HUMAN	P50461 homo sapien
24	196.5	12.0	192	1 CSR2_HUMAN	Q16527 homo sapien
25	196.5	12.0	365	1 LMCI_HUMAN	Q9NZU5 homo sapien
26	195.5	12.0	192	1 CSR2_RAT	Q62908 rattus norv
27	195	11.9	193	1 CSR2_CHICK	Q50460 gallus gall
28	195	11.9	193	1 CSR2_COTJA	Q05158 coturnix co
29	195	11.9	194	1 CSR3_MOUSE	P50462 mus musculus
30	194	11.9	194	1 CSR3_RAT	P50463 rattus norv
31	194	11.9	365	1 LMCI_MOUSE	Q8VEE1 mus musculus
32	182	11.1	192	1 CYSR_RAT	P48751 rattus norv
33	182	11.1	898	1 Z071_XENLA	P18751 xenopus lae

#### RESULT 1

ID	SLI2_MOUSE	STANDARD;	PRT;	289 AA.
AC	Q9R059; Q9JLP5; Q9WUH3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Skeletal muscle LIM-protein 2 (SLIM 2) (Four and a half LIM domains protein 3) (FHL-3).			
DE	FHL3.			
GN	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RX	MEDLINE=99160848; PubMed=10049693;			
RA	Morgan M.J., Madgwick A.J.A.;			
RT	"The LIM proteins FHL1 and FHL3 are expressed differently in skeletal muscle.";			
RL	Biochem. Biophys. Res. Commun. 255:245-250(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RX	MEDLINE=20368180; PubMed=10906474;			
RA	Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;			
RT	"Expression patterns of FHL/SLIM family members suggest important functional roles in skeletal muscle and cardiovascular system.";			
RL	Mech. Dev. 95:259-265(2000).			
RN	[3]			
RP	SEQUENCE OF 11-228 FROM N.A.			
RA	Li H.Y., Lee S.M.Y., Tsui S.K.W., Chan K.K., Kotaka M., Chim S.S.C.,			
RT	Lee C.Y., Fung K.P., Waye M.M.Y.;			
RL	"The cloning, sequencing and characterization of a mouse FHL3, which contains four and a half LIM domains.";			
CC	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
CC	-1- TISSUE SPECIFICITY: In the adult, expressed at high levels in skeletal muscle and, to a lesser extent, in heart, lung, skin and kidney. During embryonic development, expressed ubiquitously at low levels.			
CC	-1- SIMILARITY: Contains 4 LIM zinc-binding domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AF114382; AAD53231.1; -			
DR	EMBL; AF134772; AAD32623.1; -			
DR	EMBL; AF149826; AAF73159.1; -			
DR	HSP; P04006; 1IML.			
DR	MGI; MGI:1341092; Fhl3.			
DR	InterPro; IPR001781; LIM.			
DR	Pfam; PF00412; LIM; 4.			

#### ALIGNMENTS

34	179.5	11.0	3075	1	LMAL_HUMAN
35	178	10.9	192	1	CYSR_HUMAN
36	177	10.8	191	1	CYSR_CHICK
37	174.5	10.7	642	1	LK2_CHICK
38	170.5	10.4	606	1	LRG6_XENLA
39	169.5	10.4	1017	1	LRG1_YEAST
40	167	10.2	208	1	CRP2_HUMAN
41	166.5	10.2	647	1	LK1_RAT
42	166	10.2	208	1	CRP2_RAT
43	163.5	10.0	398	1	LHX3_BRARE
44	163	10.0	367	1	LHX4_HUMAN
45	163	10.0	367	1	LHX4_MOUSE

P25391	homo sapien
P21291	homo sapien
P32965	gallus gall
P53666	gallus gall
P18733	xenopus lae
P35688	saccharomyc
P52943	homo sapien
P53669	rattus norv
P36201	rattus norv
Q30421	brachydanio
Q96992	homo sapien
P53776	mus musculu

DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 3.  
 DR PROSITE; PS00023; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.  
 FT ZN\_FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 275 LIM 4.  
 FT CONFLICT 77 77 D -> G (IN REF. 1).  
 FT CONFLICT 91 91 E -> D (IN REF. 1).  
 FT CONFLICT 95 95 T -> S (IN REF. 1).  
 FT CONFLICT 137 137 A -> G (IN REF. 1 AND 3).  
 FT CONFLICT 157 159 KFA -> NLT (IN REF. 3).  
 FT CONFLICT 192 192 K -> Q (IN REF. 1).  
 FT CONFLICT 199 199 H -> Q (IN REF. 1 AND 3).  
 FT CONFLICT 205 205 D -> E (IN REF. 1).  
 FT CONFLICT 225 225 N -> K (IN REF. 1 AND 3).  
 FT CONFLICT 229 238 TGGSGGAEGA -> V (IN REF. 1).  
 FT CONFLICT 256 256 S -> N (IN REF. 1).  
 SQ SEQUENCE 289 AA; 31817 MW; 176DD857FCD1B9D9 CRC64;

Query Match 96.5%; Score 1578.5; DB 1; Length 289;  
 Best Local Similarity 93.4%; Pred. No. 1.8e-122;  
 Matches 270; Conservative 6; Mismatches 4; Indels 9; Gaps 1;  
 QY 1 MSESFDCAKNESLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELPEYDRH 60  
 DB 1 MSEAEDCAKNESLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELPEYDRH 60  
 QY 61 FHEGFCRCRCORSIADPEFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 DB 61 FHEGFCRCRCORSIADPEFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 QY 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180  
 DB 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPITGGSGGAEGAGL 240  
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPITGGSGGAEGAGL 240  
 QY 232 GGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280  
 DB 241 GGGKYVSFEDRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 289

RESULT 2  
 SLI2\_HUMAN  
 ID SLI2\_HUMAN STANDARD; PRT; 280 AA.  
 AC Q13643.  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Skeletal muscle LIM-protein 2 (SLIM 2) (Four and a half LIM domains protein 3) (FHL-3).  
 DE FHL3 OR SLIM2.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=99160848; PubMed=10049693;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal muscle.";  
 RL Biochem. Biophys. Res. Commun. 255:245-250(1999).  
 RN [2]  
 RP PRELIMINARY SEQUENCE OF 16-280 FROM N.A.  
 RC TISSUE=Skeletal muscle;

RX MEDLINE=96354835; PubMed=8753811;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "Slim defines a novel family of LIM-proteins expressed in skeletal muscle.";  
 RL Biochem. Biophys. Res. Commun. 225:632-638(1996).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN SKELETAL MUSCLE.  
 CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U60116; AAC04466.2; -.  
 DR PIR; T09504; T09504.  
 DR HSSP; P32965; ICTL.  
 DR Gene; HGNC:3704; FHL3.  
 DR MIM; 602790; -.  
 DR GO; GO:0007517; P:muscle development; TAS.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS00023; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.  
 FT ZN\_FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 275 LIM 4.  
 SQ SEQUENCE 280 AA; 31291 MW; 21AFF0950045A3DC CRC64;  
 Query Match 96.0%; Score 1569; DB 1; Length 280;  
 Best Local Similarity 95.7%; Pred. No. 1e-121;  
 Matches 268; Conservative 4; Mismatches 8; Indels 0; Gaps 0;  
 QY 1 MSESFDCAKNESLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELPEYDRH 60  
 DB 1 MSESFDCAKNESLYGRKVIQTDSPGVCPCVDNTFANTCAECQQLIGHDSRELPEYDRH 60  
 QY 61 FHEGFCRCRCORSIADPEFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 DB 61 FHEGFCRCRCORSIADPEFTQDSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGGQ 120  
 QY 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180  
 DB 121 TWHEHCFLCSCGEQPLGSRFVDPKGAHYCPVCYENKFAPRCARCKTLTGGVTVYRDQP 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPITGGVTVYRDQP 240  
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPITGGVTVYRDQP 240  
 QY 241 DRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280  
 DB 241 DRHHNCFSCARCSTSLVGQGFVDPDQVLCQGCQSQAGP 280  
 RESULT 3  
 SLI3\_RAT  
 ID SLI3\_RAT STANDARD; PRT; 279 AA.  
 AC O35115.  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)  
 DE (Four and a half LIM domains protein 2) (FHL-2).  
 GN FHL2 OR SLIM3 OR DRAL.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Histar; TISSUE=Brain;  
 RA Tanahashi H.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB008571; BAA23357.1; -  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS00023; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; zinc; Zinc-finger.  
 FT ZN\_FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 275 LIM 4.  
 SQ SEQUENCE 279 AA; 32086 MW; 9A9D8E5935034173 CRC64;  
 Query Match 61.0%; Score 998; DB 1; Length 279;  
 Best Local Similarity 55.6%; Pred. No. 5.7e-75;  
 Matches 154; Conservative 55; Mismatches 68; Indels 0; Gaps 0;  
 QY 1 MSSFDFCAKNSLYGRKVIQTDSPGYCPYDNTFANTCAEQQLIGHDSRELFDHR 60  
 DB 1 MTERFDCCHCNESLYGKVIKKEENPHCVACFEELYANTCECGTPIGCDKLSYKDRH 60  
 QY 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGG 120  
 DB 61 WHEGCFHSCSGSLVDPKFAAKEEQLLCTDCYSNEYSKQCEKKTIMPGRMEYKGS 120  
 QY 121 TWHEHCFCLSGCQPLQAGQFTSRDEPDYCVACFGLFAPKCSCKRPVIGLGGKYSFE 180  
 DB 121 SWHETCFRCQCOPIGTFKSFIPKFNQFCVPCYEKQYALQVCQCKPITGGVTRDQP 180  
 QY 181 WHRECLVCTGCTPLAQGQFTSRDEPDYCVACFGLFAPKCSCKRPVIGLGGKYSFE 240  
 DB 181 WHRECFVCTACKQLSQRTARDERPCLTFCFLYAKKACGCTNPISGLGCTKYSFE 240  
 QY 241 DRWHHNCFCARCSTSLVGQFVPGDQVLCQCSQ 277  
 DB 241 ERQWHDNCFNCCKCSLSVGRGFLTERDDILCPDCGK 277  
 RESULT 4  
 SLI3\_MOUSE STANDARD; PRT; 279 AA.  
 ID 070433; P97448;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)  
 DE (Four and a half LIM domains protein 2) (FHL-2).  
 GN FHL2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RA Chan K.K., Tsui S.K.W., Lee C.Y., Fung K.P., Waye M.M.Y.;  
 RT "The cloning, sequencing and characterization of a mouse FHL2, which  
 RT contains four and a half LIM domains.";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Skeletal muscle;  
 RX MEDLINE-99160848; PubMed-10049693;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal  
 RT muscle.";  
 RL Biochem. Biophys. Res. Commun. 255:245-250(1999).  
 RN [3]  
 RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE-20368180; PubMed-10906474;  
 RA Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;  
 RT "Expression patterns of FHL/SLIM family members suggest important  
 RT functional roles in skeletal muscle and cardiovascular system.";  
 RL Mech. Dev. 95:259-265(2000).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Starzinski-Powitz A., Martin B., Eckerdt F.;  
 RT "Isolation of the mouse homolog mDRAL from skeletal muscle derived  
 RT myoblasts.";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in heart but also detectable  
 CC in brain and skeletal muscle.  
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF055889; AAC12770.1; -  
 DR EMBL; U77040; AAB19211.2; -  
 DR EMBL; AF114381; AAD53230.1; -  
 DR EMBL; AF153340; AAD34170.1; -  
 DR MGD; MGI:1338762; Fhl2.  
 DR GO; GO:0005634; C:nucleus; IDA.  
 DR GO; GO:0005515; P:protein binding activity; IPI.  
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; IDA.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS00023; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; zinc; Zinc-finger.  
 FT ZN\_FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 275 LIM 4.  
 SQ SEQUENCE 279 AA; 32072 MW; 6D8CBC4B4424BFF2 CRC64;  
 Query Match 60.6%; Score 991; DB 1; Length 279;  
 Best Local Similarity 54.9%; Pred. No. 2.1e-74;  
 Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;  
 QY 1 MSSFDFCAKNSLYGRKVIQTDSPGYCPYDNTFANTCAEQQLIGHDSRELFDHR 60  
 DB 1 MTERFDCCHCNESLYGKVIKKEENPHCVACFEELYANTCECGTPIGCDKLSYKDRH 60  
 QY 61 FHGCFRCRCQSLADEPTCODSELLCNDYCSAFSSQCSACGETVMPGSRKLEYGG 120  
 DB 61 WHEGCFHSCSGSLVDPKFAAKEEQLLCTDCYSNEYSKQCEKKTIMPGRMEYKGS 120

QY 121 TWHEHCFCLSGCEQPLGSRSFVDPKGAHYVPCYENKFPAPCARCSKTLTGQGVYTRDQP 180  
 DB 121 SWHETCFTQRCQOPIGTGKSFIPKENQNFVPCYEKQVALQCVCKPITTTGGVYTRDQP 180  
 QY 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCRPIVGLGGKYVSFE 240  
 DB 181 WHKECFVCTACKQLSGORFTARDFFPYCLTFCDLAYAKKACGCTNPISGLGGTKYISFE 240  
 QY 241 DRWHHNCFSARCSTSLVGGVFGVDPGQVLCOGCSQ 277  
 DB 241 ERQWHDNCFNCKKSLSLVGRGFLUTERDILCPDCGK 277

RESULT 5  
 SLI3 HUMAN  
 ID SLI3 HUMAN STANDARD; PRT; 279 AA.  
 AC Q14192; Q13229; Q13644; Q9P294;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Skeletal muscle LIM-protein 3 (SLIM 3) (LIM-domain protein DRAL)  
 GN (Four and a half LIM domains protein 2) (FHL-2).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=97294674; PubMed=9150430;  
 RA Genini M., Schwalbe P., Scholl F.A., Remppis A., Mattei M.-G.,  
 RA Schaefer B.W.;  
 RT "Subtractive cloning and characterization of DRAL, a novel LIM-domain  
 RT protein down-regulated in rhabdomyosarcoma";  
 RL DNA Cell Biol. 16:433-442(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=98248917; PubMed=9573400;  
 RA Chan K.K., Tsui S.K.W., Lee S.M.Y., Luk S.C.W., Liew C.C., Fung K.P.,  
 RA Waye M.M.Y., Lee C.Y.;  
 RT "Molecular cloning and characterization of FHL2, a novel LIM domain  
 RT protein preferentially expressed in human heart";  
 RL Gene 210:345-350(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leukocyte;  
 RX MEDLINE=20458693; PubMed=11001931;  
 RA Tanahashi H., Tabira T.;  
 RT "Alzheimer's disease-associated presenilin 2 interacts with DRAL, an  
 RT LIM-domain protein";  
 RL Hum. Mol. Genet. 9:2281-2289(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEvan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 127-279 FROM N.A.  
 RC TISSUE=Heart muscle;  
 RX MEDLINE=96354835; PubMed=8753811;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "Slim defines a novel family of LIM-proteins expressed in skeletal  
 RT muscle";  
 RL Biochem. Biophys. Res. Commun. 225:632-638(1996).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED ONLY IN SKELETAL MUSCLE.  
 CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L42176; AAC5333.1; -;  
 DR EMBL; U29332; AAC52073.1; -;  
 DR EMBL; AB038794; BAA92253.1; JOINED.  
 DR EMBL; AB038991; BAA92253.1; JOINED.  
 DR EMBL; AB038992; BAA92253.1; JOINED.  
 DR EMBL; AB038792; BAA92253.1; JOINED.  
 DR EMBL; AB038793; BAA92253.1; JOINED.  
 DR EMBL; BC014397; AAH14397.1; -;  
 DR EMBL; U60117; AAC50794.1; -;  
 DR PIR; JC6565; JC6565.  
 DR Genew; HGNC:3703; FHL2.  
 DR MIM; 602633; -;  
 DR GO; GO:0005634; C:nucleus; TAS.  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; Zinc; Zinc-finger.  
 FT ZN.FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 275 LIM 4.  
 FT CONFLICT 167 167 M -> G (IN REF. 1).  
 FT CONFLICT 167 167 M -> K (IN REF. 3 AND 4).  
 SQ SEQUENCE 279 AA; 32196 MW; DBDC77C9BD6C4BBA CRC64;  
 Query Match 59.2%; Score 968; DB 1; Length 279;  
 Best Local Similarity 53.4%; Pred. NO. 1.6e-72;  
 Matches 148; Conservative 59; Mismatches 70; Indels 0; Gaps 0;  
 Qy 1 MSESFDCAKCNESLYGRKYIOTDSGPCYPCVDNTFANTCAEQCLIGHDSRELFEYDRH 60  
 Db 1 MTERFDCHCNESLFGKKYILRESYPCVCFELFANTCECGKPGICDCDLKSYKDRH 60  
 Qy 61 FHEGCFRCRCQSLADEPFTCCQDSLELNCDCYCSAFSSQCSAGCTVMPGSRKLEYGG 120  
 Db 61 WHEACFHCSCRNLSLVDPKFAAKEDQLICTDCSYNESSKQCECKKTIMPTRKMEYKGS 120  
 Qy 121 TWHEHCFCLSCCEQPLGSRSFVDPKGAHYVPCYENKFPAPCARCSKTLTGQGVYTRDQP 180  
 Db 121 SWHETCFTCHRCQOPIGTGKSFIPKENQNFVPCYEKQVALQCVCKPITTTGGVYTRDQP 180  
 Qy 181 WHRECLVTCGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCRPIVGLGGKYVSFE 240  
 Db 181 WHKECFVCTACKQLSGORFTARDFFPYCLTFCDLAYAKKACGCTNPISGLGGTKYISFE 240

QY 241 DRWHNHCSCARCSTSLVGGQFVDPDQVLCQGCSCQ 277  
 Db 241 ERQWHDNCFNCKKCSLSLVGRGFLTERDDILCPDCGK 277

## RESULT 6

SLID\_HUMAN STANDARD; PRT; 280 AA.  
 AC Q13642; Q13230; Q13645;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Skeletal muscle LIM-protein 1 (SLIM 1) (SLIM) (Four and a half LIM  
 GN domains protein 1) (FHL-1).  
 GN FHL1 OR SLIM1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=96354835; PubMed=8753811;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "Slim defines a novel family of LIM-proteins expressed in skeletal  
 muscle.";  
 RL Biochem. Biophys. Res. Commun. 225:632-638(1996).  
 RN [2]  
 RP REVISIONS.  
 RA Morgan M.J.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=98382585; PubMed=9714789;  
 RA Lee S.M.Y., Tsui S.K.W., Chan K.K., Garcia-Barcelo M., Waye M.M.Y.,  
 RA Fung K.P., Liew C.C., Lee C.Y.;  
 RT "Chromosomal mapping, tissue distribution and cDNA sequence of  
 RT four-and-a-half LIM domain protein 1 (FHL1).";  
 RL Gene 216:163-170(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Chapman J.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uslin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP SEQUENCE OF 106-255 FROM N.A.  
 RC TISSUE=Muscle;  
 RX MEDLINE=95352104; PubMed=7626119;  
 RA Morgan M.J., Madgwick A.J.A., Charlestone B., Pell J.M., Loughna P.T.;

RT "The developmental regulation of a novel muscle LIM-protein.";  
 RL Biochem. Biophys. Res. Commun. 212:840-846(1995).  
 CC -1- FUNCTION: MAY HAVE AN INVOLVEMENT IN MUSCLE DEVELOPMENT OR  
 CC HYPERTROPHY.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE AND TO A  
 CC LESSER EXTENT IN PLACENTA, OVARY, PROSTATE, TESTIS, SMALL  
 CC INTESTINE, COLON AND SPLEEN. NOT EXPRESSED IN BRAIN, LUNG, LIVER,  
 CC KIDNEY, PANCREAS, THYMUS AND PERIPHERAL BLOOD LEUKOCYTES.  
 CC -1- DEVELOPMENTAL STAGE: ELEVATED LEVELS DURING POSTNATAL MUSCLE  
 CC GROWTH.  
 CC -1- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U60115; AAC52021.1; -  
 DR EMBL; U29538; AAC35421.1; -  
 DR EMBL; AL078638; CAC18881.1; -  
 DR EMBL; BC010998; AAH10998.1; -  
 DR EMBL; U60118; AAC50795.1; -  
 DR HSSP; P32965; 1CTL.  
 DR Genew; HGNC:3702; PHL1.  
 DR MIM; 300163; -  
 DR GO; GO:0016049; P:cell growth; NAS.  
 DR GO; GO:0007517; P:muscle development; NAS.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS00033; LIM\_DOMAIN\_2; 4.  
 KW Repeat; LIM domain; Metal-binding; Zinc; Developmental protein;  
 KW Differentiation; Zinc-finger.  
 FT ZN.FING 7 31 GATA-LIKE (POTENTIAL).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 276 LIM 4.  
 FT CONFLICT 73 73 H -> Q (IN REF. 1).  
 FT CONFLICT 98 98 S -> F (IN REF. 1).  
 FT CONFLICT 158 158 F -> L (IN REF. 1 AND 6).  
 SQ SEQUENCE 280 AA; 31895 MW; 2FC873D70E62834D CRC64;  
 Query Match 46.6%; Score 762.5; DB 1; Length 280;  
 Best Local Similarity 44.2%; Pred. No. 1.le-55;  
 Matches 123; Conservative 102; Mismatches 102; Indels 1; Gaps 1;  
 QY 1 MSSSPFCACNESLYGRKYTOTDSGPYCPYDNTFANTCAEQQLIGHDSRELFYEDRH 60  
 Db 1 MAEFPCDYCRDPLQKKYQKDGHCCLCKFCDFCANTCVCCKPKTGADSKVEHYKNRF 60  
 QY 61 FHGGFCRCRCQSLADEPFTCDSELLCNDYCSAFSSOCACGETVMPGSRKLEYGGQ 120  
 Db 61 WHDTCFCAKCLHPLANETFAVDNKILCNKCTREDSPKCKGCFRAIVAGDQNVYKGT 120  
 QY 121 TWHEHCFCLSCGCEQPLGSRFSVPDKGAHYCPVCYENKFAFPCARCSTKLTGGVYTRDQP 180  
 Db 121 VHKDCFTCSNCKQVIGTSGFFPKGEDFYCVTCHETKFAKHCKVCNKAITSGGITYDQP 180  
 QY 181 WHRECLVCTCQTPLAGOQFTSRDEDPYCVACGFGELFAPKCSCKRPIVGLG-GGKYVSF 239  
 Db 181 WHADCFVCVTCCKKLQGRFQTAVEDQYVCVDCYKFNFAKCKAGCKNPITGFGKSSVYAY 240  
 QY 240 EDRHWHNHCFCSCARCTSLVGGQFVDPDQVLCQGCSCQ 277  
 Db 241 EGQSHWDYCFPHCKKCSVNLANKRPFVHQEQVCPDCAK 278

## RESULT 7

SL1L\_MOUSE STANDARD; PRT; 280 AA.  
 AC P97447: O55181; O8K318;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Skeletal muscle LIM-protein 1 (SLIM 1) (Four and a half LIM domains  
 DE protein 1) (FHL-1) (KyoT) (RBP associated molecule 14-1) (RAM14-1).  
 GN FHL1.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=99160848; PubMed=10049693;  
 RA Morgan M.J., Madgwick A.J.A.;  
 RT "The LIM proteins FHL1 and FHL3 are expressed differently in skeletal  
 RT muscle.";  
 RL Biochem. Biophys. Res. Commun. 255:245-250(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND ALTERNATIVE PRODUCTS.  
 RX MEDLINE=98078719; PubMed=9418910;  
 RA Taniguchi Y., Furukawa T., Tun T., Han H., Honjo T.;  
 RT "LIM protein KyoT2 negatively regulates transcription by association  
 RT with the RBP-J DNA-binding protein.";  
 RL Mol. Cell. Biol. 18:644-654(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.  
 RX MEDLINE=20368180; PubMed=10906474;  
 RA Chu P.-H., Ruiz-Lozano P., Zhou Q., Cai C., Chen J.;  
 RT "Expression patterns of FHL/SLIM family members suggest important  
 RT functional roles in skeletal muscle and cardiovascular system.";  
 RL Mech. Dev. 95:259-265(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Colon;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: May have an involvement in muscle development or  
 CC hypertrophy. Isoform 2 binds to RBP-J and plays a negative  
 CC regulatory role in the RBP-J-mediated transcription in mammalian  
 CC systems.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (isoform 1); nuclear (isoform  
 CC 2).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1; Synonyms=KyoT1;  
 CC IsoId=p97447-1; Sequence=Displayed;  
 CC Name=2; Synonyms=KyoT2;  
 CC IsoId=p97447-2; Sequence=VSP\_003120, VSP\_003121;

CC -!- TISSUE SPECIFICITY: Isoform 1 seems to be most abundant in each  
 CC tissue and 2 much less abundant. Isoform 1 is highly expressed in  
 CC skeletal muscle and lung, and to a lesser extent in heart, brain  
 CC and kidney. Isoform 2 was found in brain, lung kidney and genital  
 CC organs.  
 CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: U77039; AAC02727.1; --  
 DR EMBL: U41739; AAC02805.1; --  
 DR EMBL: AF114380; AAD53229.1; --  
 DR EMBL: BC031120; AAH31120.1; --  
 DR EMBL: BC029024; AAH29024.1; ALT\_INIT.  
 DR HSSP: Q05158; IQLI.  
 DR MGD: MGI:1298387; Fhl1.  
 DR InterPro: IP8001781; LIM.  
 DR Pfam: PF00412; LIM; 3.  
 DR ProDom: PD000094; LIM; 4.  
 DR SMART: SM00132; LIM; 4.  
 DR PROSITE: PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE: PS00023; LIM\_DOMAIN\_2; 4.  
 KW Differentiation; Developmental protein; Repeat; LIM domain;  
 KW Zinc-finger; Metal-binding; Zinc; Alternative splicing.  
 FT ZN\_FING 7 31 GATA-LIKE (BY SIMILARITY).  
 FT DOMAIN 40 92 LIM 1.  
 FT DOMAIN 101 153 LIM 2.  
 FT DOMAIN 162 212 LIM 3.  
 FT DOMAIN 221 276 LIM 4.  
 FT VARSPLIC 168 194  
 FT  
 FT VARSPLIC 195 280  
 FT  
 FT CONFLICT 98 98 S -> F (IN REF. 1).  
 FT SEQUENCE 280 AA; 31889 MW; 76C64FB8087BA36D CRC64;  
 SQ  
 Query Match 46.5%; Score 760.5; DB 1; Length 280;  
 Best Local Similarity 43.9%; Pred. No. 1.6e-55;  
 Matches 122; Conservative 51; Mismatches 104; Indels 1; Gaps 1;  
 QY 1 MSEFDCAKNESLYGRKVIQTDSPYCVPCYDNTFANTCAEQQLIGHDSRELEYEDRH 60  
 DB 1 MSEKFDCHYCRDPLGGKKYQKGRHCLKCFKFCANTCVDCRKPISADAKEVHKRY 60  
 QY 61 FHEGFCRCRCORSLADEPFTCODSELLCNDYCSAFSSOCSAGCETVMPGSRKLEYGQ 120  
 DB 61 WHDNCFCRAKCLHPLASTFTVSKDGKILCNKATREDSPRCCKGCFKAIAGDQNVYKGT 120  
 QY 121 TWHEHCFCLSCQEQPLGSRFSVPDKGAHYCVPCYENKFAPCARCSKTLTGQGVYTRDQ 180  
 DB 121 VVHKDCFTCSNCKQVIGTGFPPKGEDFYCVTCHETKFAKCHVCNKAITSGGITYQDP 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFELFAPKCSSCRKPVIGLG-GGKYVSF 239  
 DB 181 WHAEFCFVCTCSKLAGQRFTEVDEQYVCDYKNFVAKKAGCKNKPITGFGKSSVAY 240  
 QY 240 EDRIWHHNCFCARCSSTLVGGQFVDPDGDQVLCQCSQ 277  
 DB 241 EQSWHDYCFHCKKCSVNLANKRFVHNEQVYCPDCAK 278  
 RESULT 8  
 LPAX\_HUMAN  
 ID LPAX\_HUMAN STANDARD; PRT; 386 AA.  
 AC O60711;  
 DT 30-MAY-2000 (Rel. 39, Created)



FT	DOMAIN	211	261	LIM 2.
FT	DOMAIN	320	320	LIM 3.
FT	DOMAIN	329	379	LIM 4.
SQ	SEQUENCE	386 AA;	43332 MW;	C8D2BE61FAB11F3A CRC64;

Query Match 19.1%; Score 313; DB 1; Length 386;  
Best Local Similarity 32.8%; Pred. No. 9.le-19;  
Matches 62; Conservative 26; Mismatches 83; Indels 18; Gaps 4

QY	101	CSACGETVMPGSRKLEYGGQTWHHEHCLSCGEQPLGSRSFVPDKGAHYCVPCYENKEAP	160
Db	152	CASCOKPI--AGKVIHALGOSWHPHFVCTHCKEIGSPFFERSGLAYCPNDYHQLEFSP	209
QY	161	RCARCSKTLTQGGVYTRDQPHRECLVCTGQOTPLAGQQFTSRDEDPYCVACFGELFAPK	220
Db	210	RCAYCAAPILDKVLTAHQIWHPEHFFCSHCGEVGAEGFHEKDKPKYCRKDFLAMPSPK	269
QY	221	CSSCKRPIVLGGGKYVSFEDRHHNCFSCARCSSTSL-VGGQFVPGDGQ-----	269
Db	270	CGGCNRPVL----ENYLSAMDVWHPCEFCVGCDCFTSFSTGTFELDGRPFCELHYHRR	325
QY	270	-VLCQGCQ 277	
Db	326	GTLCHGCGQ 334	

RESULT 9

PINC_HUMAN	STANDARD;	PRT;	325 AA.
ID	PINC_HUMAN	STANDARD;	PRT;
AC	P48059; Q9BS44;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	PINCH protein (particularly interesting new Cys-His protein) (LIM and		
DE	senescent cell antigen-like domains 1).		
GN	LIMS1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal liver;		
RX	MEDLINE=94296379; PubMed=7517666;		
RT	Rearden A.;		
RT	"A new LIM protein containing an autoepitope homologous to 'senescent		
RL	cell antigen";		
RL	Biochem. Biophys. Res. Commun. 201:1124-1131(1994).		
[2]			
RP	REVISIONS TO C-TERMINUS.		
RP	Rearden A.;		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBDJ databases.		
[3]			
RP	SEQUENCE FROM N.A.		
RC	TISSUP=Bone marrow;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Hsieh F.,		
RA	Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
RA	Blakesley R.W., Truchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Groomwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzyzanski M.I., Skalski U., Smallus D.E.,		
RA	Schmerer A., Schein J.E., Jones S.J.M., Marra M.A.;		



RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ihegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
CC -1- FUNCTION: PLAYS A ROLE IN CELL DIFFERENTIATION LATE IN MYOGENESIS.  
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSION IS RESTRICTED TO THE  
CC SOMATIC, VICERAL, AND PHARYNGEAL MUSCLES. WITHIN THE SOMATIC  
CC MUSCULATURE, MLP48B IS LOCALIZED AT THE ENDS OF MUSCLES FIBERS AT  
CC THE POINT OF ATTACHMENT TO THE EPIDERMIS. THERE IS NO EXPRESSION  
CC IN CARDIAC MESODERM OR IN FAT BODY.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS BIPHASIC, PEAKING LATE IN  
CC EMBRYOGENESIS (16-24 H EMBRYOS) AND DURING THE LARVAL TO PUPAL  
CC TRANSITION, WHEN THE MUSCULATURE IS DIFFERENTIATING. FOUND IN  
CC DEVELOPING MUSCLES OF THE VISCERAL AND SOMATIC MESODERM SUBSEQUENT  
CC TO THE FORMATION OF THE MUSCLE PRECURSOR CELLS. DECREASED LEVELS  
CC ARE STILL DETECTABLE IN ADULTS.  
CC -1- SIMILARITY: TO THE VERTEBRATE CYSTEINE-RICH PROTEINS.  
CC -1- SIMILARITY: Contains 5 LIM zinc-binding domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X91245; CAA62627.1; -;  
DR EMBL; AF090832; AAC61591.1; -;  
DR EMBL; AE003672; AAF54063.1; -;  
DR HSSP; P32965; 1CTL.  
DR FlyBase; FBgn0014863; MLP84B.  
DR InterPro; IPR001781; LIM.  
DR Pfam; PF00412; LIM; 5.  
DR ProDom; PD000094; LIM; 5.  
DR SMART; SM00132; LIM; 5.  
DR PROSITE; PS00478; LIM DOMAIN 1; 5.  
DR PROSITE; PS00023; LIM DOMAIN 2; 5.  
DR Nuclear protein; Repeat: LIM domain; Metal-binding; Zinc; Myogenesis;  
KW Developmental protein; Differentiation.  
FT DOMAIN 12 63  
FT DOMAIN 65 80  
FT DOMAIN 66 71  
FT DOMAIN 120 172  
FT DOMAIN 175 180  
FT DOMAIN 178 189  
FT DOMAIN 222 274  
FT DOMAIN 276 291  
FT DOMAIN 325 377  
FT DOMAIN 379 390  
FT DOMAIN 421 473  
FT DOMAIN 475 490  
SQ SEQUENCE 495 AA; 53525 MW; 2E559B9178E54C0E CRC64;

Query Match 18.7%; Score 305; DB 1; Length 495;  
Best Local Similarity 21.6%; Pred. No. 51e-18;  
Matches 97; Conservative 40; Mismatches 121; Indels 192; Gaps 20;  
QY 5 FDCAKNESLYGRKYIQTDSGGPYCVPCYDNTF-----ANT----- 39  
DB 37 FKCGMCKNSLDSTNCTEHERELYCKTCHGRKFGPKGYGFGTGLSDMNGSQFLRENGD 96  
QY 40 -----CABCOQLIGHDSRELFFYEDRHFHEGCFRCRCQBSLA 76  
DB 97 VPSVRNGARLEPRAIARAPEGEGCPKGGYV-YAAQMLARGRSHWKECPKCGTKRKL- 154  
QY 77 DEPTTCQ--DSELLCNDICYCSAFSSQ----- 100  
DB 155 DSILCEAPDKNIYCKGCAKFKGPKGYGVGGGALQDCYAHDDGAPAIRAIDVDKI 214  
QY 101 -----CSACGETVMPGSRKLEYGGTWHFHCFLSCGCGPQLGS--RSVPDKGAIYCV 151  
DB 215 QARPGEGCPKCGGVVYAAEQKLS-KGREWHKCFNCKDKHTLDSINASDGPDRDV-YCR 272  
QY 152 PCYENKFAPR-----SSCK-----RPVIG--LGGKYVSF-----EDRHHNCF 249  
DB 273 TCYCKWGPYGYGACSGFLQDTGLTDEQISANRPYNPDTTSIKARDGEGCPRC- 328  
QY 170 TOGGVTY-----RDQPHRECLVCTGCTPL-AGQOFTSRDEDPYCVACFGELFAPKC 221  
DB 329 --GGAVFAARQQLSKGKWHKCYNCADCHRDLSVLACDGPDGDHCRACYGKLFPGKG 386  
QY 222 -----SSCK-----RPVIG--LGGKYVSF-----EDRHHNCF 249  
DB 387 FGYGHAFTLVSTSGESTIQPDGRPLAGPKTSGGCPKGFVFAAEQMIKTRTHKRCF 446  
QY 250 SCARCSTSL----VGQGFVDPGDVQLCGC 275  
DB 447 YCSDCKRLSDSTNLNDG--PDGD-IYCRAC 473  
RESULT 11  
PAXI\_HUMAN STANDARD; PRT; 591 AA.  
AC P49023; O14970; O14971; O60360;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Paxillin.  
GN PAX.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RX MEDLINE=95197488; PubMed=7534286;  
RA Salgia R., Li J.-L., Lo S.H., Brunkhorst B., Kansas G.S.,  
RA Sobhany E.S., Sun Y., Pisick E., Hallek M., Ernst T., Tantravahi R.,  
RA Chen L.B., Griffin J.D.;  
RT "Molecular cloning of human paxillin, a focal adhesion protein  
RT phosphorylated by P210BCR/ABL."  
RL J. Biol. Chem. 270:5039-5047(1995).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA).  
RA Yamagata K., Oda N., Furuta H., Vaxillaire M., Southam L., Boriraj V.,  
RA Chen X., Oda Y., Takeda J., Yamada S., Nishigori H., Lebeau M.M.,  
RA Lathrop M., Cox R.D., Bell G.I.;  
RT "Transcription map of the 5cM region surrounding the hepatocyte  
RT nuclear factor-1a/MODY3 gene on chromosome 12."  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).  
RX TISSUE=Placenta;  
RL MEDLINE=97207310; PubMed=9054445;  
RA Mazaki Y., Hashimoto S., Sabe H.;  
RT "Monocyte cells and cancer cells express novel paxillin isoforms with

RT different binding properties to focal adhesion proteins.";

RL J. Biol. Chem. 272:7437-7444(1997).

RN [4]

RP SEQUENCE FROM N.A.

RA Bradshaw H., Ozersky P.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [5]

RP PHOSPHORYLATION OF TYR-31; TYR-118 AND TYR-181.

RX MEDLINE=21634701; PubMed=11774284;

RA Iwasaki T., Nakata A., Mukai M., Shinkai K., Yano H., Sabe H.,

RA Schaefer E., Tatsuta M., Tsujimura T., Terada N., Kakishita E.,

RA Akedo H.;

RT "Involvement of phosphorylation of Tyr-31 and Tyr-118 of paxillin in

RT M1 cancer cell migration.";

RL Int. J. Cancer 97:330-335(2002).

CC -!- FUNCTION: CYTOSKELETAL PROTEIN INVOLVED IN ACTIN-MEMBRANE

CC ATTACHMENT AT SITES OF CELL ADHESION TO THE EXTRACELLULAR MATRIX

CC (FOCAL ADHESION).

CC -!- SUBUNIT: Binds in vitro to vinculin as well as to the SH3 domain

CC of c-SRC and, when tyrosine phosphorylated, to the SH2 domain of

CC V-CRK. Isoform beta binds to focal adhesion kinase but weakly to

CC vinculin. Isoform gamma binds to vinculin but only weakly to focal

CC adhesion kinase

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=3;

CC Name=Beta;

CC IsoId=P49023-1; Sequence=Displayed;

CC IsoId=P49023-2; Sequence=VSP\_003114;

CC Name=Gamma;

CC IsoId=P49023-3; Sequence=VSP\_003115;

CC -!- PTM: Phosphorylated on tyrosine residues during Integrin-mediated

CC cell adhesion, embryonic development, fibroblast transformation

CC and following stimulation of cells by mitogens.

CC -!- SIMILARITY: Contains 4 LIM zinc-binding domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL; U14588; AAC50104.1; -.

DR EMBL; U87946; AAD00648.1; -.

DR EMBL; U87941; AAD00648.1; JOINED.

DR EMBL; U87942; AAD00648.1; JOINED.

DR EMBL; U87943; AAD00648.1; JOINED.

DR EMBL; U87944; AAD00648.1; JOINED.

DR EMBL; U87945; AAD00648.1; JOINED.

DR EMBL; D86862; BAA18997.1; -.

DR EMBL; D86863; BAA18998.1; -.

DR EMBL; AC004263; AAC05175.1; -.

DR PIR; A55933; A55933.

DR PDB; 1KKY; 06-FEB-02.

DR PDB; 1KL0; 06-FEB-02.

DR Genew; HGNC:9718; PXN.

DR MIM; 602505; -.

DR GO; GO:0005875; C:microtubule associated complex; TAS.

DR GO; GO:0005515; E:protein binding activity; IPI.

DR GO; GO:0007172; P:signal complex formation; TAS.

DR GO; GO:0007165; P:signal transduction; TAS.

DR InterPro; IPR001781; LIM.

DR InterPro; IPR001904; Paxillin.

DR Pfam; PF00412; LIM; 4.

DR Pfam; PF03535; Paxillin; 1.

DR PRINTS; PR00832; PAXILLIN.

DR ProDom; PD000094; LIM; 4.

DR SMART; SM00132; LIM; 4.

DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.

DR PROSITE; PS00023; LIM\_DOMAIN\_2; 4.

CC Cytoskeleton; Phosphorylation; LIM domain; Repeat; Metal-binding;

KW

KW Zinc; Alternative splicing; 3D-structure.

FT DOMAIN 46 53

FT PRO-RICH.

FT DOMAIN 358 408

FT LIM 1.

FT DOMAIN 417 467

FT LIM 2.

FT DOMAIN 476 526

FT LIM 3.

FT DOMAIN 535 585

FT LIM 4.

FT MOD\_RES 31 31

FT PHOSPHORYLATION.

FT MOD\_RES 118 118

FT PHOSPHORYLATION.

FT MOD\_RES 181 181

FT PHOSPHORYLATION.

FT VARSPLIC 278 311

FT Missing (in isoform Alpha).

FT VARSPLIC 278 311

FT /FTID-VSP\_003114.

FT IODLEORADGERCWAAGWPRDGRSPGQDEGG -> GSW

FT PLEEVLLVLSISSVQEGEKYPHPCAAHRTSLRSPDQPP

FT PCPO (in isoform Gamma).

FT /FTID-VSP\_003115.

FT CONFLICT 73 73 G -> S (IN REF. 4).

FT SEQUENCE 591 AA; 64533 MW; OFFDC07047E96636 CRC64;

Query Match 18.6%; Score 304.5; DB 1; Length 591;

Best Local Similarity 27.5%; Pred. No. 6.7e-18;

Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;

QY 40 CAECQQLIGHDSRELFEYEDRHFEHCEGFCRCRCRQCRSLADEPFTTCQDSSELLCNCYCSAFSS 99

DB | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |

DB 358 CGACKKPIA--GQVVTAMGKTWHPEHFCVTHQCEIGSRNFFERDQGPYCEKDYHNLFSF 415

QY 100 CCSAGGETVMPGSRKLEYGGQTWHEHCFLCSCQEPGLGSRSFVPDKGAHYCVPCYENKFA 159

DB | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |

DB 416 RCYICNGPIL--DKVVTALDRWHPEHFFCAOCGGAFFEGEGHEKDKRAYCKRDFDMPA 473

QY 160 PRCARCKTLQGGVTVYRDQPMHRECLVCTGQOTPLAQOQFTSRDEDPYCVACFGELFAP 219

DB | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |

DB 474 PKCGGCARAILENYISALNTLWHPECFVCRECTFPVNGSFFEHGQGPYCEVHYHRRGS 533

QY 220 KCSSCKRPIVGLGGKGYVSFEDRHHHNCFCSCARCTSLVGQGFYDPDQDVLQCGC 275

DB | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | |

DB 534 LCSGCKPKIT---GRCITAMAKRPHFHFVCAFLKQLNKGTGFKQNDKPYCQNC 585

RESULT 12

PAXL\_CHICK

ID PAXL\_CHICK STANDARD; PRT; 559 AA.

AC P49024;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Paxillin.

GN PXN.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95197488; PubMed=7534286;

RA Saigia R., Li J.-L., Lo S.H., Brunkhorst B., Kansas G.S.,

RA Sobhany E.S., Sun Y., Pisick E., Hallek M., Ernst T., Tantravahi R.,

RA Chen L.B., Griffin J.D.;

RT "Molecular cloning of human paxillin, a focal adhesion protein

RT phosphorylated by p210BCR/ABL.";

RL J. Biol. Chem. 270:5039-5047(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95051116; PubMed=7525621;

RA Turner C.E., Miller J.T.;

RT "Primary sequence of paxillin contains putative SH2 and SH3 domain

RT binding motifs and multiple LIM domains: identification of a vinculin

RT and p215Fak-binding region.";

RL J. Cell Sci. 107:1583-1591(1994).

RN [3]

RP PHOSPHORYLATION.

RX MEDLINE=95340539; PubMed=7615549;

RA Bellis S.L., Miller J.T., Turner C.E.;  
RT "Characterization of tyrosine phosphorylation of paxillin in vitro by  
focal adhesion kinase.";  
RL J. Biol. Chem. 270:17437-17441(1995).  
CC -I- FUNCTION: CYTOSKELETAL PROTEIN INVOLVED IN ACTIN-MEMBRANE  
CC ATTACHMENT AT SITES OF CELL ADHESION TO THE EXTRACELLULAR MATRIX  
CC (FOCAL ADHESION). BINDS IN VITRO TO VINCULIN AS WELL AS TO THE SH3  
CC DOMAIN OF C-SRC AND, WHEN TYROSINE PHOSPHORYLATED, TO THE SH2  
CC DOMAIN OF V-CRK.  
CC -I- PTM: Phosphorylated on tyrosine residues during integrin-mediated  
CC cell adhesion, embryonic development, fibroblast transformation  
CC and following stimulation of cells by mitogens.  
CC -I- SIMILARITY: Contains 4 LIM zinc-binding domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U14589; AAC59665.1; -;  
DR EMBL: L30095; AAC38018.1; -;  
DR PIR: B55933; B55933.  
DR HSSP: P04006; LIML.  
DR InterPro: IPR001781; LIM.  
DR InterPro: IPR001904; Paxillin.  
DR Pfam: PF00412; LIM; 4.  
DR Pfam: PF03535; Paxillin; 1.  
DR PRINTS: P00832; PAXILLIN.  
DR ProDom: PD000094; LIM; 4.  
DR SMART: SM00132; LIM; 4.  
DR PROSITE: PS00478; LIM\_DOMAIN\_1; 4.  
DR PROSITE: PS00023; LIM\_DOMAIN\_2; 4.  
DR Cytoskeleton; Phosphorylation; LIM domain; Repeat; Metal-binding;  
KW Zinc.  
FT DOMAIN 46 53 PRO-RICH.  
FT DOMAIN 326 376 LIM 1.  
FT DOMAIN 385 435 LIM 2.  
FT DOMAIN 444 494 LIM 3.  
FT DOMAIN 503 553 LIM 4.  
FT MOD\_RES 31 31 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 118 118 PHOSPHORYLATION (BY FAK1).  
SQ SEQUENCE 559 AA; 61242 MW; 6450270D90B2DE84 CRC64;  
Query Match 18.6%; Score 303.5; DB 1; Length 559;  
Best Local Similarity 27.5%; Pred. No. 7.6e-18;  
Matches 65; Conservative 35; Mismatches 128; Indels 8; Gaps 3;  
QY 40 CAECQQLIGHDSRELVEYDRHFEGCFRCRCORSLADEPFTQDSLELLCNDYCSAFSS 99  
Db 326 CGACKKPIA--GOVVTAMGKTWHPHFVTHCQEEIGSRNFFRFDQCPYCEKDYHNLFS 383  
QY 100 QCSACGETVMPGSRKLEYGGQVTHWHEHCFHSCGCEQPLGSRSFVPDKGAHYCVCYENKFA 159  
Db 384 RCYICNGPIL--DKVVYALDRTWHPHFHFAQCGVFFGPGFEGPKGKAYCRKDYDMFA 441  
QY 160 PRCARSKTLTQGGVYTRQDPWHRCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAP 219  
Db 442 PKGGCARAILENYISALNTLWHPFCVCRECTPTFINGSFHFDGQPYCEVHYHRRGS 501  
QY 220 KCSCKRPVGLGGKGVYVSFEDRHHNCFSCARCSTSLVGQGVFDPDQVLCQG 275  
Db 502 LCSGCKQKPT---GRCIITAMGKFFHPEHVCAPFLKQLNKGFTFKEQNDKPYCQNC 553  
RESULT 13  
ID YP96\_CAEEL STANDARD; PRT; 256 AA.  
AC Q09476;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical 28.9 kDa protein C28H8.6 in chromosome III.  
GN C28H8.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N2.  
RC STRAIN-Bristol N2.  
RA Miller N., Waterston R.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: Contains 4 LIM zinc-binding domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U20861; AAA62294.1; -;  
DR PIR: E88469; E88469.  
DR HSSP: P32965; ICTL.  
DR WormPep: C28H8.6; CE01828.  
DR InterPro: IPR001781; LIM.  
DR Pfam: PF00412; LIM; 4.  
DR ProDom: PD000094; LIM; 4.  
DR SMART: SM00132; LIM; 4.  
DR PROSITE: PS00478; LIM\_DOMAIN\_1; 4.  
DR PROSITE: PS00023; LIM\_DOMAIN\_2; 4.  
DR Hypothetical protein; LIM domain; Metal-binding; Zinc; Repeat.  
FT DOMAIN 19 69 LIM 1.  
FT DOMAIN 78 128 LIM 2.  
FT DOMAIN 137 170 LIM 3.  
FT DOMAIN 196 246 LIM 4.  
SQ SEQUENCE 256 AA; 28901 MW; 8A6A70DB736B8225 CRC64;  
Query Match 18.4%; Score 300.5; DB 1; Length 256;  
Best Local Similarity 30.6%; Pred. No. 6.6e-18;  
Matches 57; Conservative 27; Mismatches 87; Indels 15; Gaps 3;  
QY 101 CSACGETVMPGSRKLEYGGQVTHWHEHCFHSCGCEQPLGSRSFVPDKGAHYCVCYENKFA 160  
Db 19 CAACGKPII--GOVVTALGKWHPEHYTCCECAELGQRPFFERNGRAFCEDYHNGFSP 76  
QY 161 RCARSKTLTQGGVYTRQDPWHRCLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAP 220  
Db 77 KCQCGHRAITDRCVSMVKNKFNHIECTFCACENQPFQEDGFHEKNGQTYCKRDFRLFAPK 136  
QY 221 CSSCKRPVGLGGKGVYVSFEDRHHNCFSCARCSTSLVGQGVFDPDQVLCQG----- 274  
Db 137 CNGCQKPT---SNFITALGTHWHPDFCVQCHGVSNFNGASFHEHNGAPLCERHYESR 192  
QY 275 ---CSQ 277  
Db 193 GSICSQ 198  
RESULT 14  
ID UN97\_CAEEL STANDARD; PRT; 348 AA.  
AC P50464;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE LIM protein unc-97.  
GN UNC-97 OR F14D12.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;



```
DR SMART; SM00132; LIM; 3.
DR PROSITE; PS00478; LIM_DOMAIN_1; 2.
DR PROSITE; PS0023; LIM_DOMAIN_2; 3.
KW LIM domain; Metal-binding; Zinc; Repeat.
FT DOMAIN 184 249 LIM 1.
FT DOMAIN 250 309 LIM 2.
FT DOMAIN 310 373 LIM 3.
FT DOMAIN 150 153 POLY-GLU.
FT DOMAIN 513 522 POLY-HIS.
FT DOMAIN 544 557 POLY-SER.
FT CONFLICT 185 185 I -> S (IN REF. 3).
FT CONFLICT 370 509 MISSING (IN REF. 3).
SQ SEQUENCE 615 AA; 68609 MW; CD024365C072B052 CRC64;

Query Match 16.2%; Score 265; DB 1; Length 615;
Best Local Similarity 29.9%; Pred. No. 1.2e-14;
Matches 55; Conservative 31; Mismatches 88; Indels 10; Gaps 6;

QY 101 CSACGETVMPGSRKLEYG---GQWHEHCFLCGCEQPLGSRSFVPDKGAHYCVPCYEN 156
Db ||: | | : | | ||: ||: | | : | | :
186 CEEGKQIGGGDIATVAFSRAGLGACWHPQCFVCTTCQELLVDLIYFYHVKYVCGRRHAE 245
QY 157 KFAPRCARCKSLTQGGVYRD-QPWHRECLVCTGCTPLAQQTFRDEDPYCVACFGE 215
Db ||| | : : | : | : | | : | | : | | :
246 CLRPCQACDEIFISPECTEAEGRHWMDHFCCFECEASLGQRYVMQSRPHCCACYEY 305
QY 216 LFAPKCSCKRPIVGLGGKYVSFEDRHH--HNCFCARCSTSLVCGGFVDPGDQVLC- 272
Db | | | | | | | | : : : : ||| ||| ||: ||: ||: ||: ||: ||:
QY 273 QGCS 276
Db : ||
364 RACS 367
```

Search completed: August 22, 2003, 19:58:08  
Job time : 26 secs





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2003, 19:54:18 ; Search time 98 seconds  
(without alignments)  
737.293 Million cell updates/sec

Title: US-09-786-135-1

Perfect score: 1635

Sequence: 1 MSESFDCAKNESLYGRKYI.....QGVPDQVLCQGCQAGP 280

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvrius.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1635	100.0	280	4 Q9BVA2	Q9bva2 homo sapien
2	1615	98.8	280	4 Q9P100	Q9p100 homo sapien
3	1007	61.6	172	4 Q96C98	Q96c98 homo sapien
4	991	60.6	279	11 Q96C99	Q96c99 mus musculus
5	934	57.1	339	5 Q9VVB7	Q9vvb7 drosophila
6	934	57.1	339	5 Q8MVZ5	Q8mvz5 drosophila
7	934	57.1	558	5 Q81Q03	Q81q03 drosophila
8	934	57.1	558	5 Q9VVB5	Q9vvb5 drosophila
9	922	56.4	525	5 Q17832	Q17832 caenorhabdi
10	866	53.0	284	11 Q9WTX7	Q9wtx7 mus musculus
11	865	52.9	284	4 Q9NTP9	Q9ntp9 homo sapien
12	861	52.7	284	4 Q9NQU2	Q9ncu2 homo sapien
13	855	52.3	284	4 Q8WU21	Q8wv21 homo sapien
14	846	51.7	291	5 Q76491	Q76491 branchiosto
15	765.5	46.8	296	6 Q9GUV4	Q9gvv4 sus scrofa
16	760.5	46.5	280	11 Q9WU44	Q9wuh4 rattus sp.

17	732.5	44.8	280	4	Q9Y630	Q9y630 homo sapien
18	730	44.6	454	5	O17833	O17833 caenorhabdi
19	721	44.1	279	11	Q9DAH0	Q9dah0 mus musculu
20	721	44.1	279	11	Q8CDC8	Q8cdc8 mus musculu
21	718	43.9	279	11	O70403	O70403 mus musculu
22	665	40.7	339	6	Q9GL05	Q9gl05 sus scrofa
23	662	40.5	323	4	Q95212	Q95212 homo sapien
24	662	40.5	323	4	Q9UKZ8	Q9ukz8 homo sapien
25	622	38.0	245	5	Q9VVB8	Q9vvb8 drosophila
26	517.5	31.7	188	6	Q8WU41	Q8wuk1 sus scrofa
27	465	28.4	194	4	Q9NZ40	Q9nz40 homo sapien
28	458	28.0	210	11	Q8K3G3	Q8k3g3 rattus norv
29	334	20.4	779	5	Q8INN7	Q8inn7 drosophila
30	331.5	20.3	784	5	Q8INN6	Q8inn6 drosophila
31	329.5	20.2	178	5	Q9VLE3	Q9vle3 drosophila
32	328.5	20.1	765	5	Q9SV19	Q9sv19 drosophila
33	328	20.1	740	5	Q9VH91	Q9vh91 drosophila
34	328	20.1	768	5	Q9VH92	Q9vh92 drosophila
35	328	20.1	806	5	Q9SV20	Q9sv20 drosophila
36	327	20.0	197	5	Q966T5	Q966t5 drosophila
37	325.5	19.9	178	5	Q8TOV8	Q8tov8 drosophila
38	324	19.8	255	5	Q8INW6	Q8inw6 drosophila
39	324	19.8	556	5	Q9GSE0	Q9gse0 drosophila
40	324	19.8	556	5	Q9VIX3	Q9vix3 drosophila
41	324	19.8	557	5	Q8MQX3	Q8mqx3 drosophila
42	324	19.8	557	5	Q8INW7	Q8inw7 drosophila
43	324	19.8	581	5	Q9VIX2	Q9vix2 drosophila
44	324	19.8	581	5	Q9BPO9	Q9bpq9 drosophila
45	314	19.2	569	5	Q8MML5	Q8mml5 dictyostel

#### ALIGNMENTS

#### RESULT 1

ID Q9BVA2 PRELIMINARY; PRT; 280 AA.  
AC Q9BVA2;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Similar to four and a half LIM domains 3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Muscle;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.  
DR EMBL; BC001351; AAH01351.1; -;  
DR EMBL; BC011697; AAH11697.1; -;  
DR HSSP; P32965; ICTL.  
DR InterPro; IPR001781; LIM.  
DR Pfam; PF00412; LIM; 4.  
DR ProDom; PD000094; LIM; 4.  
DR SMART; SM00132; LIM; 4.  
DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
DR PROSITE; PS50023; LIM\_DOMAIN\_2; 4.  
KW LIM domain; Metal-binding; zinc.  
SQ SEQUENCE 280 AA; 31192 MW; DIA037C260370DFD CRC64;

Query Match 100.0%; Score 1635; DB 4; Length 280;  
Best Local Similarity 100.0%; Pred. No. 1.4e-161;  
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSESFDCAKNESLYGRKYIQTDSGPYVPCYDNTFANTCAEQQLIGHDSRELFFEDRH 60  
|||||  
DB 1 MSESFDCAKNESLYGRKYIQTDSGPYVPCYDNTFANTCAEQQLIGHDSRELFFEDRH 60

QY 61 FHEGCFRCRCRQSLADBPFTQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGG 120  
 DB 61 FHEGCFRCRCRQSLADBPFTQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGG 120  
 QY 121 TWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180  
 DB 121 TWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240  
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240  
 QY 241 DRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280  
 DB 241 DRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280

## RESULT 2

Q9P100 PRELIMINARY; PRT; 280 AA.  
 AC Q9P100;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE LIM-only protein FHL3.  
 GN FHL3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RA Lee S.M.Y., Tsui S.K.W., Chan K.K., Kotaka M., Li H.Y., Chim S.S.C.,  
 RA Waye M.M.Y., Fung K.P., Lee C.Y.;  
 RT "Chromosomal mapping of a skeletal muscle specific LIM-Only protein  
 RT FHL3 to the distal end of the short arm of human chromosome 1.";  
 RL Somat. Cell Mol. Genet. 0:0-0(1999).  
 CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 DR EMBL; AF133732; AAF61376.1; -.  
 DR HSP; P32965; 1CTL.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.  
 DR SMART; SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS0023; LIM\_DOMAIN\_2; 4.  
 KW LIM domain; Metal-binding; Zinc.  
 SQ SEQUENCE 280 AA; 31282 MW; 95E437D966F7A6E7 CRC64;

Query Match 98.8%; Score 1615; DB 4; Length 280;  
 Best Local Similarity 98.9%; Pred. No. 1.7e-159;  
 Matches 277; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MSESFDCAKCNESLYGRKIIQDTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 60  
 DB 1 MSESFDCAKCNESLYGRKIIQDTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 60  
 QY 61 FHEGCFRCRCRQSLADBPFTQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGG 120  
 DB 61 FHEGCFRCRCRQSLADBPFTQDSSELLCNDYCSAFSSQCSAGETVMPGSRKLEYGG 120  
 QY 121 TWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180  
 DB 121 TWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKTLTGGVYTRDQ 180  
 QY 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240  
 DB 181 WHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPITVGLGGKYYVSFE 240  
 QY 241 DRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280  
 DB 241 DRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280

DB 241 DRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280  
 RESULT 3  
 Q96C98 PRELIMINARY; PRT; 172 AA.  
 ID Q96C98;  
 AC Q96C98;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 3 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 DR EMBL; BC014501; AAH14501.1; -.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 3.  
 DR ProDom; PD000094; LIM; 3.  
 DR SMART; SM00132; LIM; 3.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 2.  
 DR PROSITE; PS0023; LIM\_DOMAIN\_2; 3.  
 KW Hypothetical protein; LIM domain; Metal-binding; Zinc.  
 SQ SEQUENCE 172 AA; 18918 MW; 5918FF5E8C015BE6 CRC64;

Query Match 61.6%; Score 1007; DB 4; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-96;  
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 109 MPGSRKLEYGGQTWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKT 168  
 DB 1 MPGSRKLEYGGQTWHEHCFLCSCGEOPLGSRFVDPKGAHYCVPCYENKFAFPCARCSKT 60  
 QY 169 LTQGGVYTRDQVHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPI 228  
 DB 61 LTQGGVYTRDQVHRECLVCTGCTPLAGQOFTSRDEDPYCVACFGELFAPKCSSCKRPI 120  
 QY 229 VGLGGKYYVSPEDRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 280  
 DB 121 VGLGGKYYVSPEDRHHHNCFCARCSTSLVGGGFVDPDQVLCQCSQAGP 172

## RESULT 4

Q8VDP9 PRELIMINARY; PRT; 279 AA.  
 ID Q8VDP9;  
 AC Q8VDP9;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Four and a half LIM domains 2.  
 GN FHL2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 DR EMBL; BC021468; AAH21468.1; -.  
 DR MGD; MGI:1338762; Fhl2.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 4.  
 DR ProDom; PD000094; LIM; 4.

DR SMART, SM00132; LIM; 4.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 4.  
 KW LIM domain; Metal-binding; Zinc.  
 SQ SEQUENCE 279 AA; 32056 MW; 11CGBC4B4131EAA7 CRC64;

Query Match 60.6%; Score 991; DB 11; Length 279;  
 Best Local Similarity 54.9%; Pred. No. 9.7e-95;  
 Matches 152; Conservative 57; Mismatches 68; Indels 0; Gaps 0;

QY 1 MSESFCACAKNESLYGKRYIQTDSGPCYPCYDNTFANTCAECQQLIGHDSRELFEDRH 60  
 Db 1 MTFERDCHCHNESLYGKRYIQTDSGPCYPCYDNTFANTCAECQQLIGHDSRELFEDRH 60  
 QY 61 FHGECFRCRCQSLADEPFTQDSSELLCNDYCASFSSQCSACGTPVMPGSKLEYGGQ 120  
 Db 61 WHGECFHCSCGSLVDKPAKKEQLLCTDCYSEYSSKCECKTITMPTGRKMEYKGS 120  
 QY 121 TWHEHFCPLSCGROPGLRSFVDPDKGAYHVCYPCYENKFAPCARCCKTLFGGVYTRDQP 180  
 Db 121 SWHETCFTQCRCQOPIGTSFKPENONFCVPCYKQYALQCVQCKPKITTTGGVYTRDQP 180  
 QY 181 WHRECLVCTCCQPLAGQFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKYYVSE 240  
 Db 181 WHRECFVCTACKQLSCQRTARDEFPYLCITCDLFAKCKAGCTNPISLGGTKYLSFE 240  
 QY 241 DRHHHNCFCARCSTSLVGQGVDPGDQVLCOGCSQ 277  
 Db 241 ERQWINDCFNCKKCSLSLVGRGFLTERDILCPDCGK 277

RESULT 5  
 Q9VVVB7 PRELIMINARY; PRT; 339 AA.  
 AC Q9VVVB7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE CG32171 protein.  
 GN LMPT OR CG9959 OR GGL1914 OR CG11916 OR CG13028 OR CG32171.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN RN  
 RC SEQUENCE FROM N.A.  
 RC STRAIN-Berkeley.  
 RX MEDLINE=20196006; PubMed=107311132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A.S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleob J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Matel B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleob J., Faragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL FlyBase;  
 CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 IONS.  
 DR EMBL; AE003525; AAF49396.2;  
 DR HSP; P32965; ICTL.  
 DR FlyBase; FBgn0036672; Lmpt.  
 DR InterPro; IPR001781; LIM.  
 DR Pfam; PF00412; LIM; 5.  
 DR ProDom; PD000094; LIM; 5.  
 DR SMART; SM00132; LIM; 5.  
 DR PROSITE; PS00478; LIM\_DOMAIN\_1; 4.  
 DR PROSITE; PS50023; LIM\_DOMAIN\_2; 5.  
 KW LIM domain; Metal-binding; Zinc.  
 SQ SEQUENCE 339 AA; 38804 MW; 31883C86FF6D915B CRC64;

Query Match 57.1%; Score 934; DB 5; Length 339;  
 Best Local Similarity 52.2%; Pred. No. 9.7e-89;  
 Matches 144; Conservative 55; Mismatches 77; Indels 0; Gaps 0;

QY 2 SESFDCAKNESLYGKRYIQTDSGPCYPCYDNTFANTCAECQQLIGHDSRELFEDRH 61  
 Db 60 SGHFCWCQDSELTGQRYVTRDHPHYCIKCYENFANTCECNKIIGDSKLSYKDKHW 119  
 QY 62 HEGCFRCRCQSLADEPFTQDSSELLCNDYCASFSSQCSACGTPVMPGSKLEYGGQT 121

```
Db 120 HEACFLCFKCHLSVDKQFAGKADKIYGCNGCYDAQFASRCDCGCVFVFRAGTKMEYKTRQ 179
Qy 122 WHEHCFCLSCGCEQPLGSRFVDPDKGAHVCVPCYENKFPAPRCARCSKTLTGQGVYTRDQPW 181
Db 180 WHENCFCVCKTAIGTSFTPREQEIYACAGYEKFKATRCIKNKVITSGVYTKNFPW 239
Qy 182 HRECLVCTGCOPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKVVSPED 241
Db 240 HRECTCTHCNITLAGQFTSRDEKPYCAEFGELFAPKCTACVAKPITGIGTRFISPED 299
Qy 242 RHHHNCFCSCARCSSTLVGQGVFVDPDQVLCQGCSCQ 277
Db 300 RHHHDCFCVASCASLVGRGFTDGPDLPCDCAK 335

RESULT 6
Q8MYZ5 PRELIMINARY; PRT; 339 AA.
ID Q8MYZ5
AC Q8MYZ5
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE RE3725Op.
GN BCDNA:RE37250.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez C., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paclele J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC
CC IONS
CC EMBL; AY113460; AM29465.1; -
CC FlyBase; FBgn0063025; BCDNA:RE37250.
CC InterPro; IPR001781; LIM.
CC Pfam; PF00412; LIM; 5.
CC ProDom; PD000094; LIM; 5.
CC SMART; SM00132; LIM; 5.
CC PROSITE; PS00478; LIM_DOMAIN_1; 4.
CC PROSITE; PS50023; LIM_DOMAIN_2; 5.
CC LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 339 AA; 38834 MW; AA7D2FCA9BF1FFB8 CRC64;

Query Match 57.1%; Score 934; DB 5; Length 339;
Best Local Similarity 52.2%; Pred. No. 9.7e-89;
Matches 144; Conservative 55; Mismatches 77; Indels 0; Gaps 0;

Qy 2 SESFCAKNSLYGRKVIQTDSPGVCYPCVDNTFANTCAEQOILGHDSRELVEYDRHF 61
Db 60 SGHFCWCWDSLTQQRVIRDDHPYCIKCNVENFANTCECNKIIGIDSKDLSKDKHW 119
Qy 62 HEGCFRCRCORSLADEPFTQDSELLCNDCYSAFSSQCSACGTVMPGSKKLEYGQOT 121
Db 120 HEACFLCFKCHLSVDKQFAGKADKIYGCNGCYDAQFASRCDCGCVFVFRAGTKMEYKTRQ 179
Qy 122 WHEHCFCLSCGCEQPLGSRFVDPDKGAHVCVPCYENKFPAPRCARCSKTLTGQGVYTRDQPW 181
Db 180 WHENCFCVCKTAIGTSFTPREQEIYACAGYEKFKATRCIKNKVITSGVYTKNFPW 239
Qy 182 HRECLVCTGCOPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVIGLGGKVVSPED 241
Db 240 HRECTCTHCNITLAGQFTSRDEKPYCAEFGELFAPKCTACVAKPITGIGTRFISPED 299
```

```
Qy 242 RHHHNCFCSCARCSSTLVGQGVFVDPDQVLCQGCSCQ 277
Db 300 RHHHDCFCVASCASLVGRGFTDGPDLPCDCAK 335

RESULT 7
Q8IQO3 PRELIMINARY; PRT; 558 AA.
ID Q8IQO3
AC Q8IQO3
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG32171-PD.
GN CG32171.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Patzold M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
FT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Dou P.L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Paclele J., Paragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
```





RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J., Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J.J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weiz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RL	Nature 409:685-690(2001).
CC	-I' SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
DR	EMBL; AF083394; AAD23916.1; -;
DR	EMBL; AK006414; BAB24576.1; -;
DR	MGD; MGI:1913192; Act.
DR	InterPro; IPR000362; Fumarate_lyase.
DR	InterPro; IPR001781; LIM.
DR	Pfam; PF00412; LIM; 4.
DR	ProDom; PD000094; LIM; 4.
DR	SMART; SM00132; LIM; 4.
DR	PROSITE; PS00163; FUMARATE_LYASES; 1.
DR	PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR	PROSITE; PS50023; LIM_DOMAIN_2; 4.
KW	LIM domain; Metal-binding; Zinc.
SQ	SEQUENCE 284 AA; 32906 MW; 0537543FC0B222E4 CRC64;
Query Match 53.0%; Score 866; DB 11; Length 284;	
Best Local Similarity 47.8%; Pred. No. 9.2e-82;	
Matches 131; Conservative 57; Mismatches 86; Indels 0; Gaps 0;	
Qy	2 SESEFDCAKNESLYGRKYIOTDSGPYCVPDYNTFANTCAECQOLIGHDSRELPYERHF 61   
Dd	3 SSQDFDCQCYTSSLIGKRYVLKDDNLICISYDRIFSNYCEQCKEPIDESDKLCYKRHW 62   
Qy	62 HEGCFRCRCORSLLADEPFTCDSELCLNDCYSASFSSQSACGETVMPGSRKLEYGGQT 121   
Dd	63 HEGCFRCNCCHSLIVEKPFVAKDRLDLCTOCSNECSKCFHCXRTIMPGRKKMEFGNY 122   
Qy	122 WHEHCFLCSCGEOPLGSRSPVDKGAYHVCPCVENKFAPRCARGSKTLTGOGVTYRDQPW 181   
Dd	123 WHETCFVECHCRQPIGTGKPLISKESGNVCPCFEKEFAHYCNFKKVITSGGITFRDIQIW 182   
Qy	182 HRECLVTCGQTPLAGOOFTSRDEDPYCVACFGLFAPKSSCKRPITVGGLGGKYVSFED 241   
Dd	183 HKCEFCLSGCKRKEYELPEAFMSKDDFPFLCDLYNHLYAKKAACAATKPTITGLRGAKFIQD 242   
Qy	242 RWHHNCFSCARGSTSLVGOGFVDPDGQVLCGC 275   
Dd	243 QRWHSCEFCNGKCSVSLVGEGFLTHNMELCRKC 276   
RESULT 11	
Q9NTP9	PRELIMINARY; PRY; 284 AA.
ID	Q9NTP9
AC	Q9NTP9;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	DJ393D12.2 (Novel LIM domain protein).
DN	DJ393D12.2
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RS	Skuce C.;
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.



```

Query Match      52.7%; Score 861; DB 4; Length 284;
Best Local Similarity 48.3%; Pred. No. 3e-81;
Matches 131; Conservative 52; Mismatches 88; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
DB 6 FYCQYCTASLGGKLVKDDSPGVCYDVFNSYCECKPIESDKDFCYKDRHWHEG 65
QY 65 CFRCCRCORSIADPEFTCODSELLCNDYCYSAFSSQCSACGETVMPGSRKLEYGOTWHE 124
DB 66 CFKCTKCNHSLVERPFAAKDERLLCTECSYNECSKCFCHCKRTIMPGRKMEFKGNWHE 125
QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCRCARCSKTLTQGGVTVRDOQPHRE 184
DB 126 TCFVCENCROPIGTGKPLISKESGNYCVPCEFAHYCNFKKVTSGGITFCQDLWHE 185
QY 185 CLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYYVSFEDRHW 244
DB 186 CFLCSGCRDLCEEQFMSRDDYPFCMDCYNHLIANKVACSKPIESGLTGAKFICFQDSOW 245
QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 275
DB 246 HSECFNCGRCVSLVGKGLFQNKEIFCOKC 276

RESULT 13
Q8WW21 PRELIMINARY; PRT; 284 AA.
AC Q8WW21:
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to activator of CREM in testis (Activator of CAMP-responsive element modulator (CREM) in testis).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
DR EMBL; BC021723; AAH21723.1; -.
DR EMBL; BC029490; AAH29490.1; -.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 4.
DR ProDom; PD000094; LIM; 4.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR PROSITE; PS50023; LIM_DOMAIN_2; 4.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 284 AA; 32809 MW; AFF2435667BDF812 CRC64;

Query Match      52.3%; Score 855; DB 4; Length 284;
Best Local Similarity 48.0%; Pred. No. 1.3e-80;
Matches 130; Conservative 53; Mismatches 88; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
DB 6 FYCQYCTASLGGKLVKDDSPGVCYDVFNSYCECKPIESDKDFCYKDRHWHEG 65
QY 65 CFRCCRCORSIADPEFTCODSELLCNDYCYSAFSSQCSACGETVMPGSRKLEYGOTWHE 124
DB 66 CFKCTKCNHSLVERPFAAKDERLLCTECSYNECSKCFCHCKRTIMPGRKMEFKGNWHE 125
QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCRCARCSKTLTQGGVTVRDOQPHRE 184
DB 186 CFLCSGCRDLCEEQFMSRDDYPFCMDCYNHLIANKVACSKPIESGLTGAKFICFQDSOW 245
QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 275
DB 246 HSECFNCGRCVSLVGKGLFQNKEIFCOKC 276

RESULT 14
O76491 PRELIMINARY; PRT; 291 AA.
AC O76491:
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE LIM-domain protein.
GN AMPHIDRAL.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=9767167;
RA Schubert M., Holland N.D., Holland L.Z.;
RT "Amphioxus Amphidral encoding a LIM-domain protein: Expression in the epidermis but not in the presumptive neuroectoderm.";
RL Mech. Dev. 76:203-205(1998).
CC -1- SIMILARITY: CONTAINS 5 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC IONS.
DR EMBL; AF071773; AAC69756.1; -.
DR HSP; P32965; 1CTPL.
DR InterPro; IPR001781; LIM.
DR Pfam; PF00412; LIM; 5.
DR ProDom; PD000094; LIM; 5.
DR SMART; SM00132; LIM; 5.
DR PROSITE; PS00478; LIM_DOMAIN_1; 4.
DR PROSITE; PS50023; LIM_DOMAIN_2; 5.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 291 AA; 33430 MW; 253289F21E27E4AA CRC64;

Query Match      51.7%; Score 846; DB 5; Length 291;
Best Local Similarity 45.6%; Pred. No. 1.1e-79;
Matches 124; Conservative 63; Mismatches 85; Indels 0; Gaps 0;

QY 5 FDCAKNESLYGRKVIQTDSPGVCPCYDNTFANTCAEQQLIGHDSRELYEDRHFEH 64
DB 11 FNCFCNNNSLGHRYVNRDTHYCLKCYEKLFAPCECHGQKIGTDVLDLGFNNKHWEQ 70
QY 65 CFRCCRCORSIADPEFTCODSELLCNDYCYSAFSSQCSACGETVMPGSRKLEYGOTWHE 124
DB 71 CFNCSKCKSLVDQOFTOKSKIKYCAQCHKEFLGKDCGCHQHPDGGKMEYQGNWHE 130
QY 125 HCFLCSGCEQPLGSRFVDPDKGAHYCVPCYENKFAPCRCARCSKTLTQGGVTVRDOQPHRE 184
DB 131 KCFTCKECKKPVGKSFIAKDDKVKTCQPCYEDKVAKKCEKCRKVTSMGITYKDTPWHE 190
QY 185 CLVCTGCOTPLAGQOFTSRDEDPYCVACFGELFAPKCSCKRPVGLGGKYYVSFEDRHW 244
DB 191 CFVCTHCKKPMGSRFTSKDNNPYCINCYGDLFAKKCAKCTKPTITGLGGTKFISFNSNW 250
QY 245 HNFCSARCSTSLVGQGFVDPDGOVLQCGC 276
DB 251 HSDCFNCTGCKTSLVGKGLFQNKEIFCOKC 282

```



Search completed: August 22, 2003, 19:59:56  
Job time : 103 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 10:38:49 ; Search time 4833 Seconds  
(without alignments)  
16683.814 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 ggtccgtctgcagtcgcag.....ctctgacgtgggatgaa 1971

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*\*
- 2: gb\_hgt:\*\*
- 3: gb\_in:\*\*
- 4: gb\_cm:\*\*
- 5: gb\_ov:\*\*
- 6: gb\_pat:\*\*
- 7: gb\_ph:\*\*
- 8: gb\_pl:\*\*
- 9: gb\_pr:\*\*
- 10: gb\_ro:\*\*
- 11: gb\_sts:\*\*
- 12: gb\_sy:\*\*
- 13: gb\_un:\*\*
- 14: gb\_vi:\*\*
- 15: em\_ba:\*\*
- 16: em\_fun:\*\*
- 17: em\_hum:\*\*
- 18: em\_in:\*\*
- 19: em\_mu:\*\*
- 20: em\_or:\*\*
- 21: em\_or:\*\*
- 22: em\_ov:\*\*
- 23: em\_pat:\*\*
- 24: em\_ph:\*\*
- 25: em\_pi:\*\*
- 26: em\_ro:\*\*
- 27: em\_sts:\*\*
- 28: em\_un:\*\*
- 29: em\_vi:\*\*
- 30: em\_hgt\_hum:\*\*
- 31: em\_hgt\_inv:\*\*
- 32: em\_hgt\_other:\*\*
- 33: em\_hgt\_mus:\*\*
- 34: em\_hgt\_pln:\*\*
- 35: em\_hgt\_rod:\*\*
- 36: em\_hgt\_mam:\*\*
- 37: em\_hgt\_vrt:\*\*
- 38: em\_sy:\*\*
- 39: em\_hgt\_hum:\*\*
- 40: em\_hgt\_mus:\*\*
- 41: em\_hgt\_other:\*\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1514.2	76.8	1690	9	BC001351	BC001351 Homo sapi
2	1514.2	76.8	1691	9	BC011697	BC011697 Homo sapi
3	1473.6	74.8	1497	9	AF133732	AF133732 Homo sapi
4	1466.6	74.4	1516	9	HSU60116	U60116 Homo sapien
5	1328.2	67.4	1492	9	BC014501	BC014501 Homo sapi
c	915.2	46.4	99687	9	AL603790	AL603790 Human DNA
6	871.6	44.2	1146	10	AF134772	AF134772 Mus muscu
7	842	42.7	843	9	BT007052	BT007052 Homo sapi
8	841	42.7	843	12	BT007804	BT007804 Synthetic
9	807.6	41.0	997	4	AY277587	AY277587 Sus scrof
10	711	36.1	1147	10	AF114382	AF114382 Mus muscu
11	541.4	27.5	656	10	AF149826	AF149826 Mus muscu
c	402.2	20.4	241886	10	AC027285	AC027285 Mus muscu
13	344	17.5	840	10	AB008571	AB008571 Rattus no
14	344	16.9	876	10	AF114381	AF114381 Mus muscu
15	333.4	16.9	1300	10	AF055889	AF055889 Mus muscu
16	332.4	16.9	1433	9	HUMDRAL	L42176 Homo sapien
17	332.4	16.9	1585	6	AX015392	AX015392 Sequence
18	332.4	16.9	1892	9	BC014397	BC014397 Homo sapi
19	331.2	16.8	840	9	BT006960	BT006960 Homo sapi
20	331.2	16.8	840	12	BT007658	BT007658 Synthetic
21	330.8	16.8	1416	9	HSU29332	U29332 Homo sapien
22	326.8	16.6	263370	2	AC122957	AC122957 Rattus no
c	326.8	16.6	279467	2	AC112814	AC112814 Rattus no
24	323.8	16.4	1378	10	AF153340	AF153340 Mus muscu
25	323.8	16.4	1450	10	MMU77040	U77040 Mus muscu
26	323.8	16.4	1450	10	BC021468	BC021468 Mus muscu
27	323.8	16.4	1487	10	BC021468	BC021468 Mus muscu
c	306.2	15.5	208285	10	AC0606907	AL606907 Mouse DNA
c	306.2	15.5	219825	10	AC098886	AC098886 Mus muscu
30	296.4	15.0	318	6	AX210478	AX210478 Sequence
31	280.4	14.2	2579	4	SSC275968	AJ275968 Sus scrof
c	280.2	14.2	108296	2	AC112815	Continuation (4 of
32	276.2	14.0	1078	10	AF083394	AF083394 Mus muscu
33	276.2	14.0	1136	10	BC048473	BC048473 Mus muscu
34	265.4	13.5	2913	3	AY113460	AY113460 Drosophil
35	265.4	13.5	2916	3	AY069696	AY069696 Drosophil
36	265.4	13.5	2916	3	DME418713	AJ418713 Drosophil
37	265.4	13.3	1320	9	AK058175	AK058175 Homo sapi
38	262.4	13.3	1320	9	AK058175	AK058175 Homo sapi
39	262.4	13.3	1627	9	AF278541	AF278541 Homo sapi
40	262.4	13.3	1846	9	AK057611	AK057611 Homo sapi
41	260.8	13.2	1089	9	BC021723	BC021723 Homo sapi
42	260.8	13.2	1321	9	BC029490	BC029490 Homo sapi
43	257.2	13.0	2592	3	AF071773	AF071773 Branchios
44	256	13.0	1200	10	AF134773	AF134773 Rattus sp
45	255.4	13.0	966	9	HSU29538	U29538 Human heart

ALIGNMENTS

RESULT 1

BC001351

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1690)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

1690 bp mrna linear PRI 06-JUN-2003  
IMAGE:2964682, complete cds.

BC001351

Homo sapiens

CDNA clone MGC:8696

GI:12655006

MGC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1690)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,

	BASE COUNT	336 a	550 c	452 g	352 t
	ORIGIN				
	Query Match		76.8%	Score 1514.2;	DB 9; Length 1690;
	Best Local Similarity		98.2%	Pred. No. 0;	Mismatches 28; Indels 1; Gaps 1;
	Matches 1542;		Conservative 0;		
QY	11	CAGTCGGCAGCGCTCCGCGGACGTTCGGTCCGCCGCCCTTTGGC-CGGGCCCGCAGGGTTC	69		
Db	104	CGGTTCGCGCGCGCTGC CGCGCGCTCCCGCGCTCCCGCGCGCTTCGAGGGTTC	163		
QY	70	TCTCCCTTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGAACAAGTCCCTGTAT	129		
Db	164	TCTCCCTTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAATGAACAAGTCCCTGTAT	223		
QY	130	GGAGCAGTAGTACATCCAGACAGACAGCGGGCCCCCTACTGTGTGCCCCTGTATGACATAACC	189		
Db	224	GGAGCAGTAGTACATCCAGACAGACAGCGGGCCCCCTACTGTGTGCCCCTGTATGACATAACC	283		
QY	190	TTTGCCAACACCTGTGTGAGTGCACGAGCTATATCGGGCATGACTGCAGGGAGCTGTTC	249		
Db	284	TTTGCCAACACCTGTGTGAGTGCACGAGCTATATCGGGCATGACTGCAGGGAGCTGTTC	343		
QY	250	TATGAAGACCGCCATTTCACGAGGGGTGCTTCGGTGTGCTCCGCTGCCAGCGCTCACTA	309		
Db	344	TATGAAGACCGCCATTTCACGAGGGGTGCTTCGGTGTGCTCCGCTGCCAGCGCTCACTA	403		
QY	310	GCGGATGAACCCCTTCACCTGCCAGGACAGTGAAGTGTCTGTGCAATGACTGCTACTGCAGT	369		
Db	404	GCGGATGAACCCCTTCACCTGCCAGGACAGTGAAGTGTCTGTGCAATGACTGCTACTGCAGT	463		
QY	370	CGGTTTTCTCGCAGTGTCCGGCTTGTGGGAGACTGTCATGCTGGTCCCGGAAGCTG	429		
Db	464	CGGTTTTCTCGCAGTGTCCGGCTTGTGGGAGACTGTCATGCTGGTCCCGGAAGCTG	523		
QY	430	GAATATGGAGGCCAGACATGGCATGAGCACTGCTTCTGTGCAGTGGCTGTGAACAGCCA	489		
Db	524	GAATATGGAGGCCAGACATGGCATGAGCACTGCTTCTGTGCAGTGGCTGTGAACAGCCA	583		
QY	490	CTGGGCTCCGCTCTTTTGTGCCCGAAGGGTGTCTACTACTGCTGCTGCTCTATGAG	549		
Db	584	CTGGGCTCCGCTCTTTTGTGCCCGAAGGGTGTCTACTACTGCTGCTGCTCTATGAG	643		
QY	550	AACAGTTTGTCTCGCTGCGCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACA	609		
Db	644	AACAGTTTGTCTCGCTGCGCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACA	703		
QY	610	TACGCTGATCAGCGTGGCATCGAGAATGTCTGTGTGTATCCGGATGCCAGAGCGCCCTG	669		
Db	704	TACGCTGATCAGCGTGGCATCGAGAATGTCTGTGTGTATCCGGATGCCAGAGCGCCCTG	763		
QY	670	GCAGGACAGTAGTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTGTTGGAGAA	729		
Db	764	GCAGGACAGTAGTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTGTTGGAGAA	823		
QY	730	CTCTTTGCACCTAAAGTCAGCAGCTGCAAGCGCCCATCTGTAGTCTGCTGAGGCAAG	789		
Db	824	CTCTTTGCACCTAAAGTCAGCAGCTGCAAGCGCCCATCTGTAGTCTGCTGAGGCAAG	883		
QY	790	TATGTGTCTTTTGAAGACCGACACTGGCACCACAACCTGCTTCTCTCGCCGCCCTGCTCT	849		
Db	884	TATGTGTCTTTTGAAGACCGACACTGGCACCACAACCTGCTTCTCTCGCCGCCCTGCTCT	943		
QY	850	ACCTCCCTGTGTGGCCAGGGCTTCGTACCGATGGAGACCAAGTGTCTGCCAGGCTGT	909		
Db	944	ACCTCCCTGTGTGGCCAGGGCTTCGTACCGATGGAGACCAAGTGTCTGCCAGGCTGT	1003		
QY	910	AGCCAGGACGGCCCTTAAGCCAGGGCTTCCTGGACCCAGGCTTTCCCATACACAGGGCCCA	969		
Db	1004	AGCCAGGACGGCCCTTAAGCCAGGGCTTCCTGGACCCAGGCTTTCCCATACACAGGGCCCA	1063		
QY	970	GGACTGTGGCTTCTTTTCTAAACCACCTCTGGGACTCAGCTCCCGCCCGCAAAAAATG	1029		

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J.J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

2 (bases 1 to 1690)  
Straussberg, R.  
Direct Submission  
Submitted (12-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapbs-ref@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ruben Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;  
Web site: http://www.nisc.nih.gov/  
Contact: nisc\_mgc@nhgri.nih.gov  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 1 Row: p Column: 15  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4758373.  
Location/Qualifiers  
1..1690  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:8696 IMAGE:2964682"  
/tissue\_type="Muscle, rhabdomyosarcoma"  
/clone\_lib="NIH\_MGC\_17"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
179..1021  
/codon\_start=1  
/product="Unknown (protein for MGC:8696)"  
/protein\_id="AAH01351.1"  
/db\_xref="GI:12655007"  
/translation="MSEFDCAKNESILGRKVIQTDSGPCVPCYDNTFANTCAECQQLIGHDSRELYEDRHPHEGFCRCRCQRSLADEPTQCDSHELLCNDCYCASFSSQCSACGETVPGLSRKLGYTDQWHEHFLICSGCEQPGLGSRSFPVDGRAHYCPVCYENKFAPRCARCKMTLQDPWHREVLTCTGCTQPTPLAQQQFTSRDEPDIYCVACFGELFAPKSSCKRPPIVLGGGGRYVSFDHRHHNHCFSCARGSTSLVGQGFVDPDGDLCOGCS OAGP"

JOURNAL MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

TITLE

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK COMMENT

S

1064	Db	GGACTGTGGCTCCTCTTTTCTAAACCAACCGCTGTGGAGCTACAGCTCCCGCCGCCAAAAAATG	1123
1030	Qy	GGTCTCCTTCTGGGCTCCAGGATGTCTCCCACTCCAGCATCCCAAACTGGTACTCCC	1089
1124	Db	GGTCTCCTTCTGGGCTCCAGGATGTCTCCCACTCCAGCATCCCAAACTGGTACTCCC	1183
1090	Qy	TGACCCAGGCCCCCAATCCTGGGCTCTTACAGAGCCTCCATGAGTCAAGCCCCCTCCCA	1149
1184	Db	TGACCCAGGCCCCCAATCCTGGGCTCTTACAGAGCCTCCATGAGTCAAGCCCCCTCCCA	1243
1150	Qy	CACCTGGAGTCCAGAAATTACACCTCTCCCTCGAGTCTGGGTTCCAGACTGAGTCTCT	1209
1244	Db	CACCTGGAGTCCAGAAATTACACCTCTCCCTCGAGTCTGGGTTCCAGACTGAGTCTCT	1303
1210	Qy	CCCCAAATCAGGCGCTTAGACCCGAGCGCTTCCAAACCTGGACTCTGGAGTCTAGGCCCC	1269
1304	Db	CCCCAAATCAGGCGCTTAGACCCGAGCGCTTCCAAACCTGGACTCTGGAGTCTAGGCCCC	1363
1270	Qy	TTAAATCTAGACTTCCTTTATAGGTTTACAGTCTCCTATGGGTGCTGGGAAGTCTTGG	1329
1364	Db	TTAAATCTAGACTTCCTTTATAGGTTTACAGTCTCCTATGGGTGCTGGGAAGTCTTGG	1423
1330	Qy	AAAGTGGACTGTCTTCAGGCTTGACCTTGCCCCACCCCATCCCCCGGTTGAGGCTGTGG	1389
1424	Db	AAAGTGGACTGTCTTCAGGCTTGACCTTGCCCCACCCCATCCCCCGGTTGAGGCTGTGG	1483
1390	Qy	GGCAGCAGATCAGAGGCCCACTGATTAAGGGGCCCTAGGGTACAGGTTGTCGCCACGAG	1449
1484	Db	GGCAGCAGATCAGAGGCCCACTGATTAAGGGGCCCTAGGGTACAGGTTGTCGCCACGAG	1543
1450	Qy	TCGCCACCGAGTCTCTCATTTTATTTTCAGCTTCCATTTTGCCCATAGATGGGCAGAG	1509
1544	Db	TCGCCACCGAGTCTCTCATTTTATTTTCAGCTTCCATTTTGCCCATAGATGGGCAGAG	1603
1510	Qy	GGTGAGATTGGCTCATCCCCCTTCCAGATTCTGCAATAAAGCGGTGTGAGGAAGCAGGA	1569
1604	Db	GGTGAGATTGGCTCATCCCCCTTCCAGATTCTGCAATAAAGCGGTGTGAGGAAGCAGGA	1663
1570	Qy	AAAAAAAAAAAA 1580	
1664	Db	AAAAAAAAAAAA 1674	

RESULT 2	
BC011697	
LOCUS	1691 bp mRNA linear PRI 02-AUG-2001
DEFINITION	Homo sapiens, clone MGC:19547 IMAGE:3532981, mRNA, complete cds.
ACCESSION	BC011697
VERSION	BC011697.1 GI:15079779
KEYWORDS	MGC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1691)
TITLE	Strausberg,R.
JOURNAL	Direct Submission
REMARK	Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@nhgri.nih.gov">nisc_mgc@nhgri.nih.gov</a> Shevchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M.,

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantrop, S., Thomas, P.J., Tlonsong, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 27 Row: d Column: 9  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 4759373.

FEATURES	source	BASE COUNT
Location/Qualifiers		337 a 550 c 452 q 352 t
i. .1691		
/organism="Homo sapiens"		
/mol_type="mRNA"		
/db_xref="taxon:9606"		
/clone="MGC:19547 IMAGE:3532981"		
/tissue_type="Muscle, Rhabdomyosarcoma"		
/clone_lib="NIH_MGC_17"		
/lab_host="DH10B-R"		
/note="Vector: pOTB7"		
179. .1021		
/codon_start=1		
/product="Unknown (protein for MGC:19547)"		
/protein_id="AAH11697.1"		
/db_xref="GI:15079780"		
/translation="MSEFSDCAKNESLYGRKVIQTDSPGYPVPCYDNTFANTCAEQC QLIGHDSRELFEYDRHFHEGCFPCRCORSILADEPTCODSELLCNDCYCSAFSSQCS ACGETWPGSRKLEYGOWPHEHCFLCSCGEOPLGSRSFVPDKGAHYCYPVCYENKRAP RCARCSKTIQGGVTYRDQPWHEHRECLVTCGCTPLAGQQTSDREDDPYCVACGGEELFA PKACCSKTRPIVGLGGGKYVDRHREHWHHNCFCSCARGCSTSLVGQGFVPDGGQVILCQGCSS QAQP"		

Db	1664	AAAAAAAAAA	1674
RESULT 3	AF133732		
LOCUS	AF133732		
DEFINITION	Homo sapiens LIM-only protein FHL3 (FHL3).mRNA, complete cds.	1497 bp	linear
ACCESSION	AF133732		
VERSION	AF133732.1	GI:7381057	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1497) Lee,S.M.Y., Tsui,S.K.W., Chan,K.K., Kotaka,M., Li,H.Y., Chim,S.C., Wayne,M.Y., Fung,K.P. and Lee,C.Y.		
TITLE	Chromosomal mapping of a skeletal muscle specific LIM-Only protein FHL3 to the distal end of the short arm of human chromosome 1		
JOURNAL	Somat. Cell Mol. Genet. (1999) In press		
REFERENCE	2 (bases 1 to 1497) Lee,S.M.Y., Tsui,S.K.W., Chan,K.K., Wayne,M.M.Y., Fung,K.P., Liew,C.C. and Lee,C.Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (08-MAR-1999) Biochemistry, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China		
JOURNAL	Location/Qualifiers		
FEATURES	1..1497		
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="1" /map="lp34-p35" /tissue_type="skeletal muscle" 1..1497		
gene	/gene="FHL3"		
	/notes="four and a half LIM domains 3"		
CDS	1..843		
	/gene="FHL3"		
	/codon_start=1		
	/product="LIM-only protein FHL3"		
	/protein_id="AAF61376.1"		
	/db_xref="GI:7381058"		
	/translation="MSESTDKACNESLYGRKVIQTDSPGVPVPCVDNTFANTCAECQ OLIGHDSRELFYEDRHFEPCRCRCORSLADEPTCODSELLCNDCYCSCAFSSQCS ACGETVMPGSRKLEYGQVTHHEHCFLSCGEQPLGSRVSFPPKGARHYCPCVENKRFPA SARCSTLTQGVVTRDQPHRECLVTCGQTPPLARQPTSRDEPYPYCVACFGELFPA PKCSCKRPDIVLGGGKYVSFEDRHHWHNCFSCARCSTSLVGQGFVDPDQVLQCGQSF QAG"		
BASE COUNT	309 a	465 c	391 g 332 t
ORIGIN			
Query Match	74.8%;	Score 1473.6;	DB 9; Length 1497;
Best Local Similarity	99.1%;	Pred. No. 0;	
Matches 1482;	Conservative	0; Mismatches	14; Indels 0; Gaps 0;
QY	85	ATGAGCGAGTCATTTTGACCTGTGCAAAATGCAACGAGTCCCTGCTATGAGCGCAAGTACATC	144
DB	1	ATGAGCGAGTCATTTTGACCTGTGCAAAATGCAACGAGTCCCTGCTATGAGCGCAAGTACATC	60
QY	145	CAGACAGACAGCGGGCCCTACTGTGTGCCCTGCTATGACAATACCTTTGCGCAACACTGT	204
DB	61	CAGACAGACAGCGGGCCCTACTGTGTGCCCTGCTATGACANTACCTTTGCCACACTGT	120
QY	205	GCTGAGTGCACGAGCTTATCGGGCATGACTCGAGGGAGCTGTTCTATGAAGACCGCCAT	264
DB	121	GCTGAGTGCACGAGCTTATCGGGCATGACTCGAGGGAGCTGTTCTATGAAGACCGCCAT	180
QY	265	TTCCACGAGGGTGTCTTCGCGTCTCGCGTGCACGCGTCACTAGCCGATGACCCCTTC	324
DB	181	TTCCACGAGGGTGTCTTCGCGTCTCGCGTGCACGCGTCACTAGCCGATGACCCCTTC	240
QY	325	ACCTGCCAGGACGAGTGCCTGCTCAATGACTGCTACTGCAGTGCCTTTCTCTCCGCGAG	384



Db	782	GCTCACTCGCACC	CGCTGCTCTAACTCCCTGGTGGCCACGGCTTCGTACCGGATGGAG	841
QY	887	ACCAAGTGCTCTGCC	AGGGCTGTAGCCAGCAGCGGGCCCTAAGCAGGGCTCTCTGGACCCA	946
Db	842	ACCAAGTGCTCTGCC	AGGGCTGTAGCCAGCAGCGGGCCCTAAGCAGGGCTCTCTGGACCAA	901
QY	947	GGCTTTTCCCATACC	ACGGGCCACAGGACTGTGGCTCTCTTTTCTAAACACACCTCTGGGACTC	1006
Db	902	GGCTTTCTCTATACC	ACGGGCCACAGGACTGTGGCTCTCTTTTCTAAACACACCTCTGGGACTC	961
QY	1007	AGCTCCCCCGCCCA	AAAAAATGGGTCTCTCTTGGGCTCCAGATTGTCTCCCACTCC	1066
Db	962	AGCTCCCCCGCCCA	AAAAAATGGGTCTCTCTTGGGCTCCAGATTGTCTCCCACTCC	1021
QY	1067	AGCATCCCCAAACT	GTGTACTCCCTGACCCAGGGCCCAACTCTGGGCTCTTACAGAGCCT	1126
Db	1022	AGCATCCCCAAACT	GTGTACTCCCTGACCCAGGGCCCAACTCTGGGCTCTTACAGAGCAT	1081
QY	1127	CCATGAGTCAAGC	CCCCCTCCCCACACCTGGACTCCAGAAATTCACCCCTCTCCCTCGAGTC	1186
Db	1082	CCATGAGTCAAGC	CCCCCTCCCCACACCTGGACTCCAGAAATTCACCCCTCTCCCTCGAGTC	1141
QY	1187	TGGGTTCCCAAGACT	GTAGTCTCTCCCAATACAGGGCTCTAGACCGAGCCCTCCAAACC	1246
Db	1142	TGGGTTCCCAAGACT	GTAGTCTCTCCCAATACAGGGCTCTAGACCGAGCCCTCCAAACC	1201
QY	1247	TGGACTCTGGGACT	TTAGGCCCTTAACTAGACTTCTCTTATAGTTTCAGGTCTCC	1306
Db	1202	TGGACTCTGGGACT	TTAGGCCCTTAACTAGACTTCTCTTATAGTTTCAGGTCTCC	1261
QY	1307	TATGGGTGCTGGG	AAGTCTTGAAGTGGACTGTCTCAGGGCTTGACCTGCCCCACCCC	1366
Db	1262	TATGGGTGCTGGG	AAGTCTTGAAGTGGACTGTCTCAGGGCTTGACCTGCCCCACCCC	1321
QY	1367	ATCCCCGGGTTGAG	GCTGTGGGGCAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAG	1426
Db	1322	ATCCCCGGGTTGAG	GCTGTGGGGCAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAG	1381
QY	1427	GGTACAGGGTGCT	CCCCAGCAGCTGCCACCGAGTGCTCTCATTTATTTTTCAGCTCCA	1486
Db	1382	GGTACAGGGTGCT	CCCCAGCAGCTGCCACCGAGTGCTCTCATTTATTTTCAGCTCCA	1441
QY	1487	TTTTTGCCATPAGAT	TGGCAGAGGGGTGAGATTGGCTCATCCCCCTCCAGATTCTGCAAT	1546
Db	1442	TTTTTGCCATPAGAT	TGGCAGAGGGGTGAGATTGGCTCATCCCCCTCCAGATTCTGCAAT	1501
QY	1547	AAAGCGGTGTGAGGA	1561	
Db	1502	AAAGCGGTGTGAGGA	1516	
RESULT 5				
BC014501				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REMARK				
COMMENT				







Db 68994 CGAGTGTCTCTCATTTTATTTTTCAGTCCATTTTGTCCCATAGATGGCAGAGGGTGAGA 68935  
QY 1517 TTGGCTCATCCCTTCCAGATCTGCAATAAAGCGGTGTGAGGAAGCAGGG 1568  
|||||  
Db 68934 TTGGCTCATCCCTTCCAGATCTGCAATAAAGCGGTGTGAGGAAGCAGGG 68883  
|||||  
RESULT 7  
AF134772 1146 bp mRNA linear ROD 27-MAY-1999  
LOCUS Mus musculus LIM protein (Fhl3) mRNA, complete cds.  
DEFINITION AF134772  
ACCESSION AF134772  
VERSION AF134772.1 GI:4894846  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1146)  
The LIM proteins FHL1 and FHL3 are expressed differently in  
skeletal muscle  
Morgan, M.J. and Madgwick, A.J.  
Biochem. Biophys. Res. Commun. 255 (2), 245-250 (1999)  
JOURNAL 99160848  
MEDLINE 10049693  
PUBMED  
REFERENCE 2 (bases 1 to 1146)  
AUTHORS Morgan, M.J.  
TITLE Mouse FHL3 is expressed at high levels in adult skeletal muscle  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 1146)  
AUTHORS Morgan, M.J.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAR-1999) Orthodontics, Eastman Dental Institute, 256  
Gray's Inn Road, London WC1X 8LD, UK  
FEATURES  
source 1..1146  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
gene 1..1146  
/gene="Fhl3"  
CDS 172..1014  
/gene="Fhl3"  
/note="FHL3"  
/codon\_start=1  
/product="LIM protein"  
/protein\_id="AAB32623.1"  
/db\_xref="GI:4894847"  
/translation="MSEAFDCAKCNESLYGRKIOTDSGPYCPYCDNTFANTCABCQ  
QLTHDSRELFYEDHFHGFRCRCORSLAGEPFTCODSELLCNDYCSAFSSQCS  
ACGETVMPGSRKLEYGGTWHHCFLCSCGEOPLGSRSFVPDKGAHYCPYENKEFAP  
RCARCSKTLTGQGVTYRDQPHRECLVCTGCGTPLAGQQFTSRDEDPYCVACFEGELFA  
PKSSCKRPIVLGGGKYSVFEDRHWHNCFSCARSTSLVGQGFVPDGDQVLCQCS  
QAGP"  
misc\_feature 190..264  
/gene="Fhl3"  
/note="Region: HL-domain"  
misc\_feature 289..447  
/gene="Fhl3"  
/note="Region: LIM-domain 1"  
misc\_feature 472..630  
/gene="Fhl3"  
/note="Region: LIM-domain 2"  
misc\_feature 655..807  
/gene="Fhl3"  
/note="Region: LIM-domain 3"  
misc\_feature 832..996  
/gene="Fhl3"  
/note="Region: LIM-domain 4"  
BASE COUNT 207 a 394 c 313 g 232 t  
ORIGIN

Query Match 44.2%; Score 871.6; DB 10; Length 1146;

Best Local Similarity 92.1%; Pred. No. 1.8e-226;  
Matches 930; Conservative 0; Mismatches 79; Indels 1; Gaps 1;  
QY 13 GTCCGAGAGCTCGCCGCGCAGCTCGCTCGGCGCCCGCTTGGCGCGCCCGCAGGGTCTCT 72  
Db 99 GGCTCCGCGCCCGCCACCGCCGCGCGCCGCTCCGCGCGCGCTTCGAGGGCTCACT 158  
QY 73 -CCCTTTGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAGTCCCTGTATGG 131  
Db 159 GGGCACCAGCCACCATGAGCGAGCATTTGACTGTGCAAAATGCAACGAGTCCCTGTACGG 218  
QY 132 AGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGTGTATGCAATACCTT 191  
Db 219 CCGCAATACATCCAGACAGACAGTGGCCCTACTGCTTCCGTTCCTGTATGACAACACTT 278  
QY 192 TGCCAAACCTGTGTAGTGCAGCAGCTTATCGGCGATGATCGGAGGAGCTGTCTTA 251  
Db 279 CGCCAAACCTGTGTAGTGCAGCAGCTTATCGGCGATGATCGGAGGAGCTGTCTTA 338  
QY 252 TGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTCGGCTGCCAGCGCTCACTAGC 311  
Db 339 TGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTCGGCTGCCAGCGCTCACTAGC 398  
QY 312 CGATGAACCTTCCACCTGCCAGACAGTGTGCTGTGCAATGACTGTCTACTGCACTGC 371  
Db 399 CGGTGAACCTTCCACCTGCCAGACAGTGTGCTGTGCAATGACTGTCTACTGCACTGC 458  
QY 372 GTTTTCTCGACGTCTCCGCTTGTGGGAGAGCTGTCATCGCTGGGTCGCGGAGCTGGA 431  
Db 459 GTTTTCTCGACGTCTCCGCTTGTGGGAGAGCTGTCATCGCTGGGTCGCGGAGCTGGA 518  
QY 432 ATATGAGGCGCAGACATGGCATGAGCAGCTCTTCTGTGCGAGTGGCTGTGAACAGCCACT 491  
Db 519 ATATGAGGCGCAGACATGGCATGAGCAGCTCTTCTGTGCGAGTGGCTGTGAACAGCCACT 578  
QY 492 GGGCTCCGTTCTTTGTGCCCGACAGAGGTGTCTACTGTGCGTGCCTGTCTATGAGAA 551  
Db 579 GGGCTCCGTTCTTTGTGCCCGACAGAGGTGTCTACTGTGCGTGCCTGTCTATGAGAA 638  
QY 552 CAAGTTGTCTCTCGCTGCCCGCTGCAGCAGACAGCTGCACAGGCTGAGTGACATA 611  
Db 639 CAAGTTGTCTCTCGCTGCCCGCTGCAGCAGACAGCTGCACAGGCTGAGTGACATA 698  
QY 612 CCGTGATCAGCGCTGGCATCGAAGATGTCTGGTGTGTACCGGATGCCAGAGCCCTTGGC 671  
Db 699 CCGTGATCAGCGCTGGCATCGAAGATGTCTGGTGTGTACCGGATGCCAGAGCCCTTGGC 758  
QY 672 AGGGCAGCAGTTCACTCCCGGATGAAGATCCCTACTGTGTGGCTGTGTTTGGAGAAT 731  
Db 759 AGGGCAGCAGTTCACTCCCGGATGAAGATCCCTACTGTGTGGCTGTGTTTGGAGAAT 818  
QY 732 CTTTGACCTTAAGTGAGCAGAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCACTA 791  
Db 819 CTTTGACCTTAAGTGAGCAGAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGCACTA 878  
QY 792 TGTGTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTCGCCCGCTCTCTAC 851  
Db 879 TGTGTCTTTGAAGACCGACACTGGCACCACTGCTTCTCTCGCCCGCTCTCTAC 938  
QY 852 CTCCTGTGTGGCGCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGGCTGTAG 911  
Db 939 CTCCTGTGTGGCGCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGGCTGCAG 998  
QY 912 CCAGGAGGGCCCTTAAGCCAGGGCTCTGTGACCCAGGCTTTCATACCGAGGCCCAGG 971  
Db 999 CCAAGCAGGGCCCTTGAGTGGAGGATCTGTCTCTCTTCCCAACCAAGAGTTCCTCAAG 1058  
QY 972 ACTGTGGCTCTCTTTCTAAACCACTCTGGGAGCTCAGCTCCCGCCGCAA 1021  
Db 1059 ACTTTGGCGCTCTTTCTGAACCACTTCTGGGGCCCAACCCCTCCCTAGAA 1108

RESULT 8

```

BT007052      843 bp      mRNA      linear      PRI 13-MAY-2003
LOCUS         Homo sapiens four and a half LIM domains 3 mRNA, complete cds.
DEFINITION   BT007052.1 GI:30582942
ACCESSION    BT007052.1
VERSION      FLI_CDNA.
SOURCE       Homo sapiens (human)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 843)
AUTHORS      Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
              Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
              Phelan,M. and Farmer,A.
TITLE        Cloning of human full-length CDSs in BD Creator(TM) System Donor
              vector
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 843)
AUTHORS      Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
              Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
              Phelan,M. and Farmer,A.
TITLE        Direct Submission
JOURNAL      Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
              Circle, Palo Alto, CA 94303, USA
COMMENT      This CDS clone is a part of a collection of human full length
              expression clones generated by BD Biosciences Clontech and the
              Harvard Institute of Proteomics. Each CDS has been cloned in two
              forms: with and without stop-codon (to allow fusion with C-terminal
              tag). The CDS has been directionally cloned using BD In-Fusion(TM)
              cloning system between the SalI and HindIII sites of the pDNR-DUAL
              vector. Additional sequences in the clone: 'Acc' after SalI site
              and before 'ATG' to provide kozak consensus sequence; 'GG' after
              last codon and before HindIII site to maintain reading frame.
              Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES     Location/Qualifiers
              1..843
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="GH00305X1.0"
                 /clone_lib="BD Creator(TM) CDS Library derived from MGC
                 collection"
                 /lab_host="DH5alpha T1 resistant"
                 /note="Vector: pDNR-Dual"
                 1..843
                    /codon_start=1
                    /product="four and a half LIM domains 3"
                    /protein_id="AAP35701.1"
                    /db_xref="GI:30582943"
                    /translation="MSEFDCAKNESLYGRKYIOTDSGPYVPCYDNTFANTCAEQ
                    QLIGHDSRELFEDRHFEHFGCFRCQQRSLADEPFTQDSLELNCDCYCSAFSSQCS
                    ACGETVMPGSRKLEYGGTWHCHFLCSCGQPLGSRSFVPDKGAHYKCPVKENKFAF
                    RCARSKTTLTGQVYTRDPQWRECLVCTGCOTPLAGQOFTSRDEDPYCVACFGLFA
                    PKCSCKRPIVLGGKYSFEDRHHNHCSCARCSLSLVGQGFVDPDQVLCQGS
                    QAGP"
BASE COUNT   164 a 251 c 243 g 185 t
ORIGIN
Query Match      42.7%; Score 842; DB 9; Length 843;
Best Local Similarity 100.0%; Pred. No. 2.2e-218;
Matches 842; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      85  ATGACGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 144
      Db      1  ATGACGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 60
QY      145  CAGACAGACAGCGGCCCTACTGTGTCCTGTATGACATACCTTTGCCAACACCTGT 204
      Db      61  CAGACAGACAGCGGCCCTACTGTGTCCTGTATGACATACCTTTGCCAACACCTGT 120
QY      205  GCTGATGCCAGCAGCTTATCGGCATGACCTCGAGGAGCGTGTCTATGAAGACCGGCAT 264
      Db      121  GCTGATGCCAGCAGCTTATCGGCATGACCTCGAGGAGCGTGTCTATGAAGACCGGCAT 180

```

```

QY      265  TTCACAGAGGCGTCTTCCGCTGCTGCCAGCGCTCAGTACCGATGAACCCCTTC 324
      Db      181  TTCACAGAGGCGTCTTCCGCTGCTGCCAGCGCTCAGTACCGATGAACCCCTTC 240
QY      325  ACCTGCCAGGACAGTGAAGTCTCTGCAATGACTCTACTGCAGTGCCTTTTCCTCGAG 384
      Db      241  ACCTGCCAGGACAGTGAAGTCTCTGCAATGACTCTACTGCAGTGCCTTTTCCTCGAG 300
QY      385  TGCTCCGCTTGTGGGGAGACTGTATGCTGGTCCCGAAGCTGGAATATGGAGGCCAG 444
      Db      301  TGCTCCGCTTGTGGGGAGACTGTATGCTGGTCCCGAAGCTGGAATATGGAGGCCAG 360
QY      445  ACATGGCATGAGCACTGCTTCTGTGCAAGTGGCTGTGAACAGCCACTGGCTCCCGTCT 504
      Db      361  ACATGGCATGAGCACTGCTTCTGTGCAAGTGGCTGTGAACAGCCACTGGCTCCCGTCT 420
QY      505  TTTGTGCCCGACAGAGGTGCTCACTACTGCTGCTGCTCTATGAGAACAAAGTTGCTCCT 564
      Db      421  TTTGTGCCCGACAGAGGTGCTCACTACTGCTGCTGCTCTATGAGAACAAAGTTGCTCCT 480
QY      565  CGCTGCCGCCCTGCAGCAAGCGCTGACACAGGCTGGAGTGACATACCGTATCAGCCG 624
      Db      481  CGCTGCCGCCCTGCAGCAAGCGCTGACACAGGCTGGAGTGACATACCGTATCAGCCG 540
QY      625  TGGCATCGAGAATGTCTGGTCTGTACCGGATGCCAGACGCCCTGGCGAGGCGCAGTTC 684
      Db      541  TGGCATCGAGAATGTCTGGTCTGTACCGGATGCCAGACGCCCTGGCGAGGCGCAGTTC 600
QY      685  ACCTCCCGGATGAAGATCCCTACTGTGCTGCTGTTTGGAGAACTCTTTGCACCTAAG 744
      Db      601  ACCTCCCGGATGAAGATCCCTACTGTGCTGCTGTTTGGAGAACTCTTTGCACCTAAG 660
QY      745  TGCAGCAGCTCAAGCGCCCATCGTAGGACTCGGTGAGGCAAGTATGTCTCTTTCGA 804
      Db      661  TGCAGCAGCTCAAGCGCCCATCGTAGGACTCGGTGAGGCAAGTATGTCTCTTTCGA 720
QY      805  GACCGACACTGGCACCACTGCTTCTCTGCGCCCGCTGCTCTACTCCCTGGTGGGC 864
      Db      721  GACCGACACTGGCACCACTGCTTCTCTGCGCCCGCTGCTCTACTCCCTGGTGGGC 780
QY      865  CAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGCTGTAGCAGGAGGCC 924
      Db      781  CAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGCTGTAGCAGGAGGCC 840
QY      925  TA 926
      Db      841  TA 842

```

```

RESULT 9      BT007804      843 bp      mRNA      linear      SYN 13-MAY-2003
LOCUS         Synthetic construct Homo sapiens four and a half LIM domains 3
DEFINITION   BT007804.1 GI:30584446
ACCESSION    BT007804.1
VERSION      FLI_CDNA.
KEYWORDS     synthetic construct
SOURCE       synthetic construct
ORGANISM     artificial sequences.
REFERENCE    1 (bases 1 to 843)
AUTHORS      Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
              Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
              Phelan,M. and Farmer,A.
TITLE        Cloning of human full-length CDSs in BD Creator(TM) System Donor
              vector
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 843)
AUTHORS      Kainine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
              Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,
              Phelan,M. and Farmer,A.
TITLE        Direct Submission

```

JOURNAL

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA  
This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal-tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'gg' after last codon and before HindIII site to maintain reading frame.  
Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

location/Qualifiers  
1..843  
/organism="synthetic construct"  
/mol\_type="mRNA"  
/db\_xref="taxon:32630"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
1..>843  
/note="Mutations: 842:Stop->Leu"  
/codon\_start=1  
/transl\_table=11  
/product="Homo sapiens four and a half LIM domains 3"  
/protein\_id="AAP36476.1"  
/db\_xref="GI:30584447"  
/translation="MSEFDCAKNESLYGRKYIQDNGPYCVCYDNTFANTCAEQ  
OLIGHDSRELFEYDRHFEHFCRCRCORSLADEPFTCDSELICNDYCSAFSSQCS  
ACGTVMPGSRKLEYGGOTWHEHCFCLSGCEOPLGSRFVDPDKAHVCYVCYENKFP  
KCARCSKTLTGGVTVYRDQPHWRECLVCTGCOTPLAGQOFTSRDEDPYCVACGELE  
PKCSSKRPIVGLGGKYYVFEHRRHHNCFSCARCSLSLVGGQFVDPDQVLCQGS  
QAGP"

CDS

BASE COUNT  
ORIGIN

163 a 251 c 243 g 186 t

Query Match

Best Local Similarity 42.7%; Score 841; DB 12; Length 843;  
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 ATGACGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 144  
DB 1 ATGACGAGTCATTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCAAGTACATC 60  
QY 145 CAGACACACGCGCCCTACTGTGCGCTGTGCTGCTATGACAAATACCTTTGCCAACACCTGT 204  
DB 61 CAGACACACGCGCCCTACTGTGCGCTGTGCTGCTATGACAAATACCTTTGCCAACACCTGT 120  
QY 205 GCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGCTGTCTATGAGACCGGCAT 264  
DB 121 GCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGAGGCTGTCTATGAGACCGGCAT 180  
QY 265 TTCCACGAGGCTGCTCCGCTGTGCGCTGTGCGAGGCTCAGTACGCGATGAACCCCTTC 324  
DB 181 TTCCACGAGGCTGCTCCGCTGTGCGCTGTGCGAGGCTCAGTACGCGATGAACCCCTTC 240  
QY 325 ACCTGCCAGGACAGTGCCTCTGCAAGACTGCTACTGCACTGGGCTTTTCTCCGAG 384  
DB 241 ACCTGCCAGGACAGTGCCTCTGCAAGACTGCTACTGCACTGGGCTTTTCTCCGAG 300  
QY 385 TGCTCCGCTTGTGGGAGACTGTCATCGTGGTCCCGAGCTGGAATATGGAGCCAG 444  
DB 301 TGCTCCGCTTGTGGGAGACTGTCATCGTGGTCCCGAGCTGGAATATGGAGCCAG 360  
QY 445 ACATGGCATGAGCACTGCTTCTGTCAGTGGCTGTGAACGACCTGGGCTCCCGTCT 504  
DB 361 ACATGGCATGAGCACTGCTTCTGTCAGTGGCTGTGAACGACCTGGGCTCCCGTCT 420  
QY 505 TTTGTGCCGCAACAGGCTGCTCACTACTGCTGCCCTGCTATGAGACAAGTTTGTCTCT 564  
DB 421 TTTGTGCCGCAACAGGCTGCTCACTACTGCTGCCCTGCTATGAGACAAGTTTGTCTCT 480

QY 565 CGCTGCGCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGTGATCAGCGC 624  
DB 481 CGCTGCGCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGTGATCAGCGC 540  
QY 625 TGGCATCAGAAATGCTGGTCTGTACCGGATGCGCAGAGCCCTGGCAGGCGACGAGTTC 684  
DB 541 TGGCATCAGAAATGCTGGTCTGTACCGGATGCGCAGAGCCCTGGCAGGCGACGAGTTC 600  
QY 685 ACCCTCCGGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGCACCTAAG 744  
DB 601 ACCCTCCGGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGCACCTAAG 660  
QY 745 TGCAGCAGCTGCAAGCGCCCATCTAGTAGTCTGGTGGAGCAAGTAGTGTCTCTTTGAA 804  
DB 661 TGCAGCAGCTGCAAGCGCCCATCTAGTAGTCTGGTGGAGCAAGTAGTGTCTCTTTGAA 720  
QY 805 GACGACACTGGCACCACAACTGCTTCCTCGCCCGCTGCTTACCTCCTGGTGGGC 864  
DB 721 GACGACACTGGCACCACAACTGCTTCCTCGCCCGCTGCTTACCTCCTGGTGGGC 780  
QY 865 CAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGCTGTAGCCAGGCGGCC 924  
DB 781 CAGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGCTGTAGCCAGGCGGCC 840  
QY 925 T 925  
DB 841 T 841  
RESULT 10  
AY277587  
LOCUS AY277587  
DEFINITION Sus scrofa four and a half LIM domains 3 (PHL3) mRNA, complete cds.  
ACCESSION AY277587  
VERSION AY277587.1 GI:30526304  
KEYWORDS  
SOURCE Sus scrofa (pig)  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE  
AUTHORS Zuo,B., Xiong,Y.Z. and Deng,C.Y.  
TITLE Characterization of the porcine four and a half LIM domains 3  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 997)  
AUTHORS Zuo,B., Xiong,Y.Z. and Deng,C.Y.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2003) College of Animal Science and Veterinary  
Medicine, Huazhong Agriculture University, Agricultural Ministry  
Key Laboratory of Swine Genetics and Breeding, Loin Mountain Street  
No.1, Wuhan, Hubei 430070, P. R. China  
FEATURES  
source  
1..997  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/tissue\_type="skeletal muscle"  
/note="breed: Landrace"  
1..997  
/gene="PHL3"  
7..849  
/gene="PHL3"  
/codon\_start=1  
/product="four and a half LIM domains 3"  
/protein\_id="AAP32084.1"  
/db\_xref="GI:30526305"  
/translation="MSEFDCAKNESLYGRKYIQDNGPYCVCYDNTFANTCAEQ  
OLIGHDSRELFEYDRHFEHFCRCRCORSLADEPFTCDSELICNDYCSAFSSQCS  
ACGTVMPGSRKLEYGGOTWHEHCFCLSGCEOPLGSRFVDPDKAHVCYVCYENKFP  
KCARCSKTLTGGVTVYRDQPHWRECLVCTGCOTPLAGQOFTSRDEDPYCVACGELE  
PKCSSKRPIVGLGGKYYVFEHRRHHNCFSCARCSLSLVGGQFVDPDQVLCQGS  
QAGP"

[illegible][illegible]









Db 97900 GGCTCCTTAAGTCAAACTTCTAGGCTAGATTTTTTTAGTTTCTGCTGAGTGCCTAAG 97841  
QY 1321 AAGTCCCTTGAAGTGGACTCTTCTCAGGCTT-----GACCTGCCCCACCCCATCCCC 1372  
Db 97840 AAGTCTTCAAAATAGATTGTATACACCGGTTTGACCTCAACCCCATCCCCACTGCTCTC 97781  
QY 1373 GCGTTTGAGGCTGTGGGGGAGCAGATCAGGAGCCCACTGATAAGGGGCCCTTAGGGTACA 1432  
Db 97780 TGAGCTCGGCCCATATGATATAGCAGATTAGGGCTCTCTGTTAGAGGGTTGTAAGGTACA 97721  
QY 1433 GGGTGTGCCAGCAGGTGCCACCGAGTCTCTCTCATTTTATTCAGCTCCATTGTC 1492  
Db 97720 GGGTGTGTCCAGCCTCTGAGTAC-----CATGTTATTTTCAGCTCCATTGTC 97673  
QY 1493 CCATAGATGGCAGAGGGGTGAGATGGCTCATCCCTTCCAGATTCTGCAATAAAGCG 1552  
Db 97672 CCTGCAACAGAGAGCAG--GAAGGTGTTCACCTGTTTCCAGAGATTGTAATAAGCA 97616  
QY 1553 GTGTGAGGAAGCAGGGAAAAA 1573  
Db 97615 GTGTGAGGGGGGAAAAA 97595

## RESULT 14

AB008571  
LOCUS  
DEFINITION Rattus norvegicus mRNA for DRAL, complete cds.  
ACCESSION AB008571  
VERSION AB008571.1 GI:2605503  
KEYWORDS DRAL.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (sites)  
AUTHORS Tanahashi, H. and Tabira, T.  
TITLE Alzheimer's disease-associated presenilin 2 interacts with DRAL, an  
LIM-domain protein  
JOURNAL Hum. Mol. Genet. 9 (15), 2281-2289 (2000)  
MEDLINE 20458893  
PUBMED 11001931  
REFERENCE 2 (bases 1 to 840)  
AUTHORS Tanahashi, H.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1997), Hiroshi Tanahashi, National Institute of  
Neuroscience, NCNP, Division of Demyelinating Disease and Aging;  
4-1-1 Ogawahigashi, Kodaira, Tokyo 187-8502, Japan  
(E-mail: tanahashencnaxp.ncnp.go.jp, Tel:81-42-341-1717,  
Fax:81-42-346-1747)

## FEATURES

source  
1..840  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/tissue\_type="brain"  
1..840  
/codon\_start=1  
/product="DRAL"  
/protein\_id="BAA23357.1"  
/db\_xref="GI:2605504"  
/translations="MVERFDCHNCSNLSYKVKYLKRNPHCVACPEELVANTCEEG  
TPIGCDKDLSDRWHEGCFHSCSGSSLDKPPAAKEQLLCTDCYSNEYSSKQ  
ECKTIPGTRKMEYKSSWHETCFQCRCQCPQIGKSPFKENQNFPCVPEKQYAL  
QCVCKPITIGVTYRDQWHRCECFVCTACKKQLSQRFETARDEFPYCLTFCFDLIA  
KKCAGTNPISGLGGTKYISFEERQHNDCFNCKKCSLSLVGRGFLTERDDIILCPDGG  
KDI"  
117..279  
/note="encodes LIM-domain 1"  
301..459  
/note="encodes LIM-domain 2"  
484..636

## misc\_feature

117..279

## misc\_feature

301..459

## misc\_feature

484..636

misc\_feature /note="encodes LIM-domain 3"  
661..825  
BASE COUNT 202 a 227 c 234 g 177 t  
ORIGIN  
Query Match 17.5%; Score 344; DB 10; Length 840;  
Best Local Similarity 63.3%; Pred. No. 1.9e-82;  
Matches 527; Conservative 0; Mismatches 305; Indels 0; Gaps 0;  
QY 85 ATGAGCAGTCATTGTTGACTGTGCAAAATGCAACAGTCCCTGTATGAGCAGCAAGTACATC 144  
Db 1 ATGACTGAAGCGCTTGACTGCAACACTGCTCTCTCTATATGGAAGTACATTT 60  
QY 145 CAGACACAGCGGCCCTTACTGTGCTCCCTGTATGACATATCTTTGCAACACCTGT 204  
Db 61 CTAAGGAGGAGAACCCCGCACTGTGTGCTCTTCGAGGAGCTCTATGCCAACACCTGT 120  
QY 205 GCTGAGTGCCAGCAGCTTATCGGCGCATGACTCGAGGAGCTGTCTTATGAAGACCGCCAT 264  
Db 121 GAGGAGTGTGAACACACCATCGGCTGTGACTGCAAGGACTGTCTCTACAAGATCGCCAC 180  
QY 265 TTCCACAGGCGCTCTCCGCTGCTGCGCTGCGCAGCGCTCAGTACAGCGGATGACACCTTC 324  
Db 181 TGGCATGAAGGCTGCTTCCACTGCTCCGGTGCAGGAGCTGCTGGTGGACAAGCCCTTT 240  
QY 325 ACCTGCCAGCAGTGTGAGTCTGCAATGACTGCTACTGCACTGCGCTTTCCTCGCAG 384  
Db 241 GCGGCCAAGGAGGAGCAGCTGCTGTGCAACGACTGCTATTCACAGGACTCTCTGCTCCAG 300  
QY 385 TGCTCCGCTTGTGGGAGACTGTCTGCTGGGTCCCGGAGCTGGAATATGAGGAGCAG 444  
Db 301 TGCCAGGAGTGAAGAGACCATCATGCCAGTACCCAGAGATGGAATACAAGGCGAGC 360  
QY 445 ACATGGCATGAGCAGCTCTTCTCTGCAAGTGGCTGTGAACGACACTGGGCTCCGTTCT 504  
Db 361 AGCTGGCAGCAGAGCTGCTTCACTGCTGAGGAGCTGCCAGCAGCCCATCGGAACAGAGC 420  
QY 505 TTGTGCCCCAGACAGGCTGCTACTGCTGCTGCTGCTGCTATGAGACAGCTTGTCTCT 564  
Db 421 TTCTATCTTAAAGGAGATCAGAACTTCTGCTGCTGCTGCTATGAGACAGTATGCCCTG 480  
QY 565 CGCTGCCGCCCTGTCAGACAGCTGACACAGGCTGGAGTGACATACCGTGATCAGCG 624  
Db 481 CAGTGTGTGAGTGCAAAAAGCCAAATACACAGGAGGTGTACTTACCGGGATCAGCCC 540  
QY 625 TGGCATCGAAGTGTCTGCTGTACCGGATGCCAGAGCCCTGGCAGGCGAGCAGTTC 684  
Db 541 TGGCAGAGGAGTGTCTTGTGTGACAGCCTGCAAGAAGCAGCTATCTGGGCAACGCTTC 600  
QY 685 ACCTGCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGAGAACTCTTTGCAACCTAAG 744  
Db 601 ACAGCCCGGATGAGTGTTCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 745 TGCAGCAGCTGCAAGCGCCCATCGTAGGACTCGTAGGAGGCAAGTATGCTTCTCTTTGAA 804  
Db 661 TGTGCTGGTGCACCAACCCCATCAGTGTCTGGTGGTGGCACAAGATACATCTCTTTGAG 720  
QY 805 GACCGACACTGGGACACCAACTGCTTCTCTGCGCCGCTGCTCTACTCTCTCTGCTGCTGCTG 864  
Db 721 GAAGCCGAGTGGCACAACGACTGCTTAACTTGAAGAAGTGTCTCTCTCTGCTGCTGCTGCTG 780  
QY 865 CAGGGCTTCGTACCGGATGAGACCAAGTGTCTGCCAGGCTCTGTAGCAGG 916  
Db 781 CGAGGCTTCTCAGAGAGAGACACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832

## RESULT 15

AF114381  
LOCUS  
DEFINITION Mus musculus four and half LIM domain protein 2 (Phl2) mRNA,  
complete cds.  
ACCESSION AF114381

```

VERSION AF114381.1 GI:5825390
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 876)
AUTHORS Chu, P.H., Ruiz-Lozano, P., Zhou, Q., Cai, C. and Chen, J.
TITLE Expression patterns of FHL/SLIM family members suggest important
functional roles in skeletal muscle and cardiovascular system
JOURNAL Mech. Dev. 95 (1-2), 259-265 (2000)
MEDLINE 20368180
PUBMED 10906474
REFERENCE 2 (bases 1 to 876)
AUTHORS Chen, J.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1998) Medicine, UCSD, 9500 Gilman Drive #5024,
San Diego, CA 92093-0613C, USA
FEATURES
    source
        1..876
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            /db_xref="Fhl2"
            /gene="Fhl2"
            /codon_start=1
            /product="four and half LIM domain protein 2"
            /protein_id="AAD53230.1"
            /translation="MTERFDCHECNESLYGKRYILKEENPHCVAFEEELYANTCEECG
            TPIGDCDLVSKDRHSGCFHCSRSSLDVDPFAAKEEQLICTDYSNEYSKCO
            ECKTITMPGTRKMEYKSGSWHETCTCORCOPICTKSFIPKFNCFVPCYKQYAL
            OCVOCKKPIITGGTYIREQPHKKECFVCTACKKQLSGORFTARBEFPYCLICFCDLYA
            KRCACATNPISGLGRTKYSIFEERQWHDNCFNCKCSLSVLVGRGFLTERDDILCPDGC
            KDI"
    gene
        1..876
    CDS
        12..851
            /product="four and half LIM domain protein 2"
            /protein_id="AAD53230.1"
            /db_xref="GI:5825391"
            /translation="MTERFDCHECNESLYGKRYILKEENPHCVAFEEELYANTCEECG
            TPIGDCDLVSKDRHSGCFHCSRSSLDVDPFAAKEEQLICTDYSNEYSKCO
            ECKTITMPGTRKMEYKSGSWHETCTCORCOPICTKSFIPKFNCFVPCYKQYAL
            OCVOCKKPIITGGTYIREQPHKKECFVCTACKKQLSGORFTARBEFPYCLICFCDLYA
            KRCACATNPISGLGRTKYSIFEERQWHDNCFNCKCSLSVLVGRGFLTERDDILCPDGC
            KDI"
BASE COUNT 219 a 232 c 242 g 183 t
ORIGIN
    Query Match 16.9%; Score 333.4; DB 10; Length 876;
    Best Local Similarity 62.3%; Pred. No. 1.5e-79;
    Matches 523; Conservative 0; Mismatches 316; Indels 0; Gaps 0;
    QY 78 TGCCACCATGAGCGAGTCATTGACTGTGCAGAAATGCACAGCTCCCTGTATGGACGCAA 137
    DB 11 ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 5 TGTACGATGACTGAACGCTTTGACTGCCACACTGCATGAATCTCTGTACGGCAAGAA 64
    QY 138 GTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGTATGACATACCTTTGCCAA 197
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 65 GTACATTTAAAGGAGAGAGACCCACACTGTGTGCCCTGTCTTGTAGGAACCTATATGCAA 124
    QY 198 CACCTGTGCTGAGTCCAGCAGCTTATCGGCATGACTCGAGGGAGCTGTCTATGAAGA 257
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 125 TACCTGTGAGGAGTGTGAACACCATCGGCTGTACTGCAAGGACTTGTCTCTACAAGGA 184
    QY 258 CGCCATTTCCACGAGGCTGCTCCGCTGTGCGCTGCCAGCGCTCACTAGCCGATGA 317
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 185 TCGGCACTGGCATGAAGGCTGCTTCCACTGTCCCGTGGGGAGCTCGCTGGTGACAA 244
    QY 318 ACCCTTACCTGCCAGGACAGTGAAGCTGCTGTGCAATGACTGCTACTGAGTGCCTTTTC 377
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 245 GCCCTTTGCGGCAAGAGGAGCAGCTGCTGTGCACCGACTGCTATTCCAACGAGTACTC 304
    QY 378 CTCGAGTGTCCGCTTCTGGGAGACTGTATGCTGGTCCCGGAGCTGGAATATGG 437
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 305 GTCCAAGTGCAGGAGTGTAGAAGACCATCATCGGCTACCCCGCAGATGGAATCAA 364
    QY 438 AGGCCAGACATGGCATGACCTGCTTCTGTGAGTGGCTGTGAACCCACTGGGCTC 497
    DB 11 ||||| || || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
    DB 365 GGGCAGCAGCTGGCAGGAGACTGCTTCACTGTGAGGCTGCCAGAGCCCATTTGGAAC 424
    QY 498 CCGTTCCTTTGTGCGCCGACAAAGGGTGTCTACTGTGCTGCGCTGCTATGAGAACAAAGTT 557

```

Search completed: August 27, 2003, 13:24:08  
Job time : 4841 secs

```

Db 425 CAAGAGCTTCATACCTTAAGGAGAATCAGAACTTCTGCGTCCCTGCTATGAGAAGCAGTA 484
QY 558 TGCTCCTCGCTGCGCCCGCTGCAGCAAGACGCTGCACACAGGTGGAGTGACATACCGTGA 617
Db 485 TCCCTCTGAGTGGTGGCAGTGCAGAAAGCCTATACACAGGAGGTGTTACTTACCGGGA 544
QY 618 TCAGCCGTGGCATCGAGAATGTCTGTGTACCGGATGCCAGAGCCCTGCCAGGCA 677
Db 545 GCAGCCCTGGCACAGGAATGCTTTGTGTGCACAGCCTGCAAGAAGCAGCTATCTGGGCA 604
QY 678 GCAGTTCACTCCCGGATGAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTGC 737
Db 605 ACGCTTTCACAGCAGCGGATGAGTTTCCATACTGCTGCTGCTGCTGCTGCTGCTGCTG 664
QY 738 ACCTAAGTGCAGCAGCTGCAAGCGCCCATCTGATAGACTCGTGGAGGCAACTATGTGTC 797
Db 665 TAAGAAATGTGTGGTGCACCAACCCCATTAGTGGTGTGGGTGGCACAAGATACATCTC 724
QY 798 CTTTGAAGACCGACACTGCGCACCAACTGCTTCTCTGCGCCGCTGCTCTTACCTCCCT 857
Db 725 TTTGAGGAACGCCAATGGCACACGACTGCTTTAACTGTAGAAGTGTCTCTCTCT 784
QY 858 GGTGGCCAGGCTTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGGGTGTAGCCAGG 916
Db 785 GGTGGGAGGAGGCTTCTCTCACAGAGAGATGACATCCTCTGCTGCTGCTGCTGCTGCTG 843

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 10:37:18 ; Search time 394 seconds

(without alignments)  
13504.042 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 ggtcgtcgtcagtcgcag.....ctcttgacgtggggatgaa 1971

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_13Jun03.\*

1:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25:	/SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1971	100.0	1971	21	AAZ99798
2	332.4	16.9	1585	20	AAZ41281
3	296.4	15.0	318	22	AAH81611
4	262.4	13.3	1158	24	ABQ99467
5	255.4	13.0	1673	20	AAZ42134
6	255.4	13.0	2340	24	ABK34596
7	252.2	12.8	2254	24	ABL62451
8	252.2	12.8	2254	24	ABL66934

9	244.4	12.4	2310	22	AAH23101
10	240	12.2	2254	25	ACA56524
11	232.6	11.8	830	22	AAH01353
12	232.6	11.8	830	22	AAH33146
13	232.6	11.8	830	23	ABL98606
14	225	11.4	2042	24	ABT01067
15	225	11.4	2154	25	ABX63260
16	196.4	10.0	239	22	AAH82192
17	176.8	9.0	883	23	ABL15937
18	165	8.4	167	16	AAH21976
19	140.4	7.1	1284	22	AAF25354
20	117.6	6.0	469	21	AAH01535
21	115.8	5.9	723	24	AAH62000
22	107.4	5.4	1311	23	AAH76147
23	92.4	4.7	707	23	ABL18295
24	88.6	4.5	550	21	AAH93589
25	81.2	4.1	1593	23	ABL20585
26	81.2	4.1	1793	23	ABL19665
27	81.2	4.1	12706	23	ABL19664
28	81.2	4.1	14899	23	ABL20584
29	75.6	3.8	1486	24	AAH24709
30	75.4	3.8	1456	21	AAH87741
31	75.4	3.8	1488	20	AAH24321
32	75.4	3.8	1488	21	AAH87714
33	75.4	3.8	1575	21	AAH87742
34	75.4	3.8	1597	17	AAH12908
35	75.4	3.8	1620	21	AAH87737
36	75.4	3.8	1644	20	AAH24322
37	75.4	3.8	1644	21	AAH87715
38	75.4	3.8	1665	20	AAH24328
39	75.4	3.8	1665	21	AAH87738
40	75.4	3.8	1689	20	AAH24327
41	75.4	3.8	1689	21	AAH87729
42	75.4	3.8	1725	24	ABK83879
43	75.4	3.8	1835	21	AAH18013
44	75.2	3.8	7302	23	ABL15936
45	75	3.8	802	24	ABS51464

ALIGNMENTS

RESULT 1

AAZ99798

ID AAZ99798 standard; cDNA; 1971 BP.

XX AAZ99798;

AC AAZ99798;

XX 12-JUL-2000 (first entry)

DT cDNA encoding a human LIM domain protein homologue.

DE Human; LIM domain protein homologue; LIM domain; L protein; I protein;

KW M protein; Incyte clone 1925017; cancer; reproductive disorder;

KW autoimmune disorder; ss.

XX Homo sapiens.

OS Key

FX CDS

FT Location/Qualifiers

FT 85..927

FT /\*tag= a

FT /product= "LIM domain protein homologue"

XX WO200014231-A1.

XX 16-MAR-2000.

XX 01-SEP-1999; 99WO-US20239.

XX 03-SEP-1998; 98US-0155260.

XX (INCY-) INCYTE PHARM INC.

XX

PI Tang YT, Guegler KJ, Corley NC, Patterson C;  
XX WPI: 2000-256982/22.  
DR P-PSDB; AY84378.  
XX  
PT A purified polypeptide designated LDPH (LIM domain protein homolog),  
PT useful for treating, preventing, and diagnosing various cancers,  
PT reproductive disorders and autoimmune disorders -  
XX  
PS Claim 9; Page 53-54; 59pp; English.  
XX  
CC The present sequence encodes a human LIM domain protein homologue.  
CC The LIM domain is so named because it was first described in three  
CC proteins from *Drosophila melanogaster*, designated L, I, and M. The LIM  
CC domain is rich in cysteine. Nucleic acids encoding the LIM protein  
CC homologue were first identified in incyte clone 1925017 from a breast  
CC tissue cDNA library. The LIM polynucleotide and polypeptide are  
CC useful for treating, preventing, and diagnosing various cancers,  
CC reproductive disorders and autoimmune disorders.  
XX  
SQ Sequence 1971 BP: 399 A; 618 C; 528 G; 426 T; 0 other;  
  
Query Match 100.0%; Score 1971; DB 21; Length 1971;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1971; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GGTCCGCTGAGTCGGCAGCTCGCGGACAGCTCGTCGGCCCGCCCTGGCCCGGCC 60  
DB 1 GGTCCGCTGAGTCGGCAGCTCGCGGACAGCTCGTCGGCCCGCCCTGGCCCGGCC 60  
  
QY 61 GCAGGGTCTCCCTTGGCCACCATGAGCGAGTCATTGACTGTGCAAAATGCAACGAG 120  
DB 61 GCAGGGTCTCCCTTGGCCACCATGAGCGAGTCATTGACTGTGCAAAATGCAACGAG 120  
  
QY 121 TCCCTGTATGACCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCCCCTGCTAT 180  
DB 121 TCCCTGTATGACCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCCCCTGCTAT 180  
  
QY 181 GACAAATACCTTTGCCAACACCTGTGCTGAGTGCCAGACAGCTTATCGGGCATGACTCGAGG 240  
DB 181 GACAAATACCTTTGCCAACACCTGTGCTGAGTGCCAGACAGCTTATCGGGCATGACTCGAGG 240  
  
QY 241 GAGCTGTTCTATGAAGACCGCCATTTCCAGAGGGCTGCTTCCGCTGCTGCCGTGCCAG 300  
DB 241 GAGCTGTTCTATGAAGACCGCCATTTCCAGAGGGCTGCTTCCGCTGCTGCCGTGCCAG 300  
  
QY 301 CGCTCACTAGCCGATGAACCTTCCACCTGCCAGACAGAGTGCCTGCAATGACTGC 360  
DB 301 CGCTCACTAGCCGATGAACCTTCCACCTGCCAGACAGAGTGCCTGCAATGACTGC 360  
  
QY 361 TACTGCAGTGGTTTTCTCTCGAGTGTCTCGCTTGTGGGAGACTGTCATGCTGGGTCC 420  
DB 361 TACTGCAGTGGTTTTCTCTCGAGTGTCTCGCTTGTGGGAGACTGTCATGCTGGGTCC 420  
  
QY 421 CGGAAGCTGGAATATGAGGCGCAGATGGCATGAGCACTGCTTCTGTGCAAGTGGCTGT 480  
DB 421 CGGAAGCTGGAATATGAGGCGCAGATGGCATGAGCACTGCTTCTGTGCAAGTGGCTGT 480  
  
QY 481 GAACAGCACTGGGCTCCCGTCTTTGTCGCCAGCAAGGCTGCTACTGCTGCC 540  
DB 481 GAACAGCACTGGGCTCCCGTCTTTGTCGCCAGCAAGGCTGCTACTGCTGCC 540  
  
QY 541 TGCTATGAGAACAAATTTGCTCTCGCTGGCCCGCTGCAGCAAGAGCTGACACAGGGT 600  
DB 541 TGCTATGAGAACAAATTTGCTCTCGCTGGCCCGCTGCAGCAAGAGCTGACACAGGGT 600  
  
QY 601 GGAGTGACATACCGTGATCAGCCGTGGCATGCGAGAATGTCTGTCTGTACGGATGCCAG 660  
DB 601 GGAGTGACATACCGTGATCAGCCGTGGCATGCGAGAATGTCTGTCTGTACGGATGCCAG 660  
  
QY 661 ACGCCCTTGGCAGGCGAGCTTCCACCTGCCGGATGAAGATCCTACTGTGCGCCTGT 720  
DB 661 ACGCCCTTGGCAGGCGAGCTTCCACCTGCCGGATGAAGATCCTACTGTGCGCCTGT 720

QY 721 TTTGGAGAACTCTTTTGACACCTAAGTCAGCAGCTGCAAGCGCCCATCTAGTACTCGGT 780  
DB 721 TTTGGAGAACTCTTTTGACACCTAAGTCAGCAGCTGCAAGCGCCCATCTAGTACTCGGT 780  
  
QY 781 GGAGCAAGTATGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTCTCTCGGCC 840  
DB 781 GGAGCAAGTATGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTCTCTCGGCC 840  
  
QY 841 CGCTGCTCTACCTCCCTGGTGGCCAGGGCTTCTACCGGATGAGACCAAGTCTCTGC 900  
DB 841 CGCTGCTCTACCTCCCTGGTGGCCAGGGCTTCTACCGGATGAGACCAAGTCTCTGC 900  
  
QY 901 CAGGGCTGTAGCCAGGAGGSCCTAAGCCAGGCTCTCTGACCCAGGCTTTTCCCATACC 960  
DB 901 CAGGGCTGTAGCCAGGAGGSCCTAAGCCAGGCTCTCTGACCCAGGCTTTTCCCATACC 960  
  
QY 961 ACGGGCCAGGACTGTGGCTCTTTTCTAAACCACTCTTGGGACTCAGCTCCCCCGCA 1020  
DB 961 ACGGGCCAGGACTGTGGCTCTTTTCTAAACCACTCTTGGGACTCAGCTCCCCCGCA 1020  
  
QY 1021 AAAAAATGGTCTCTCTGGGCTCCAGGATTTGTCCTCCACATCCAGCATCCCAACT 1080  
DB 1021 AAAAAATGGTCTCTCTGGGCTCCAGGATTTGTCCTCCACATCCAGCATCCCAACT 1080  
  
QY 1081 GGTACTCCCTGACCCAGGCCCCCAATCTTGGGCTCTTACAGAGCTCCATGAGTCAAGCC 1140  
DB 1081 GGTACTCCCTGACCCAGGCCCCCAATCTTGGGCTCTTACAGAGCTCCATGAGTCAAGCC 1140  
  
QY 1141 CCCTCCCCACACTGGACTCCAGAAATTCACCTCTCCCTCGACTCTGGGTTCACAGACT 1200  
DB 1141 CCCTCCCCACACTGGACTCCAGAAATTCACCTCTCCCTCGACTCTGGGTTCACAGACT 1200  
  
QY 1201 GAGTCTCTCCCAAAATCAGGGCTTAGACCCGAGCCCTCAAACTGGGACTCTGGGACT 1260  
DB 1201 GAGTCTCTCCCAAAATCAGGGCTTAGACCCGAGCCCTCAAACTGGGACTCTGGGACT 1260  
  
QY 1261 TAGGCCCTTAAATCTAGACTTCTCTTTATAGTTTTCAGGTCTCTATGGGTGCTGGG 1320  
DB 1261 TAGGCCCTTAAATCTAGACTTCTCTTTATAGTTTTCAGGTCTCTATGGGTGCTGGG 1320  
  
QY 1321 AAGTCTTTGAAAGTGGACTGTCTCAGGCTTGACTGCTGCCACCCCATCCCGGGTGA 1380  
DB 1321 AAGTCTTTGAAAGTGGACTGTCTCAGGCTTGACTGCTGCCACCCCATCCCGGGTGA 1380  
  
QY 1381 GGCTGTGGGCGACAGATCAGGAGCCACATGATAAGGGGCCCTAGGTTACAGGTGCTG 1440  
DB 1381 GGCTGTGGGCGACAGATCAGGAGCCACATGATAAGGGGCCCTAGGTTACAGGTGCTG 1440  
  
QY 1441 CCCAGCAGGTGCGCACCGAGTGTCTTCTCAATTTTATTTAGCTCCATTTTGGCCATAGAT 1500  
DB 1441 CCCAGCAGGTGCGCACCGAGTGTCTTCTCAATTTTATTTTTCAGTCCATTTTGGCCATAGAT 1500  
  
QY 1501 GGGCAGAGGGTGGATTGGCTCATCCCCCTTCCAGATTTCTGCAATTAAGCGGTGTGAGG 1560  
DB 1501 GGGCAGAGGGTGGATTGGCTCATCCCCCTTCCAGATTTCTGCAATTAAGCGGTGTGAGG 1560  
  
QY 1561 AAGCAGGAAAAAAGGGGGCGCGGCTATTGGACTCGCGACCCGGAATAA 1620  
DB 1561 AAGCAGGAAAAAAGGGGGCGCGGCTATTGGACTCGCGACCCGGAATAA 1620  
  
QY 1621 TTTCCGCGCGGTTCCCTGGGAGGAACAGTGTTCCTTAAAGTGGCTCGTTTAAAAACC 1680  
DB 1621 TTTCCGCGCGGTTCCCTGGGAGGAACAGTGTTCCTTAAAGTGGCTCGTTTAAAAACC 1680  
  
QY 1681 TGGGGAAAAACAGGGGTAAAAACCTTCTCCCGGGGGAGAAATTTTCTCCCGGCC 1740  
DB 1681 TGGGGAAAAACAGGGGTAAAAACCTTCTCCCGGGGGAGAAATTTTCTCCCGGCC 1740  
  
QY 1741 TCAATATTTCCCGCACCAACCAATCATGAGCGGGCCCGGGGAAAGCCATATAAGTGT 1800  
DB 1741 TCAATATTTCCCGCACCAACCAATCATGAGCGGGCCCGGGGAAAGCCATATAAGTGT 1800



PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX Rosenthal A, Hinzmann B, Schaefer R, Zuber J, Tchernitsa O;  
PI Grips M, Hellriegel M, Schmitz A, Sers C;  
XX WPI; 2001-483415/52.  
XX  
XX Nucleic acids differentially expressed between tumor and normal cells,  
PT useful for diagnosis or therapy of tumors and for screening active  
PT agents  
XX  
XX Disclosure; Page 362; 579pp; German.  
XX  
XX This invention describes a nucleic acid (I) with differential expression  
CC between tumour and normal cells and which has cytostatic activity. (I)  
CC work as modulators of Ras activity by inducing expression of tumour  
CC suppressor genes. (I), and polypeptides encoded by them, are useful as  
CC targets for diagnosis or therapy and in screening to determine the  
CC effects of an active compound (potential pharmaceutical) on a cell line,  
CC particularly for diagnosis and treatment of tumors, especially by  
CC modulating expression of (I) (by gene therapy, antisense RNA or ribozyme  
CC methods) or by modulating the amount and/or location of (I)-encoded  
CC polypeptides (by administration of the polypeptide or its activator,  
CC antibody (optionally as a conjugate) or inhibitor). The method allows  
CC identification of many Class II tumour suppressor genes (i.e. genes that  
CC are not primary targets for tumour-initiating mutations).  
CC AAH81492-AAH82376 represent the human and rat derived nucleic acid  
CC fragments described in the method of the invention.  
XX  
XX Sequence 318 BP; 53 A; 97 C; 89 G; 69 T; 0 other;

Query Match 15.0%; Score 296.4; DB 22; Length 318;  
Best Local Similarity 98.0%; Pred. No. 7.3e-76;  
Matches 300; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 130 GGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTATGACAAATACC 189  
DB 2 GGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGCTGCTATGACAAATACC 61  
QY 190 TTGTCACACACCTGTGCTGAGTGCCAGCAGCTTATCGGCGATGACTCGAGGAGCTGTC 249  
DB 62 TTGTCACACACCTGTGCTGAGTGCCAGCAGCTTATCGGCGATGACTCGAGGAGCTGTC 121  
QY 250 TATGAAGACCGCATTTCCAGAGGGCTGCTTCCGCTGCTGCGCTGCCAGCGCTCACTA 309  
DB 122 TATGAAGACCGCATTTCCAGAGGGCTGCTTCCGCTGCTGCGCTGCCAGCGCTCACTA 181  
QY 310 GCCGATGACCTTACCTGCCAGCAGTACGCTGCTGCTGCAATGACTGCTACTGCACT 369  
DB 182 GCCGATGACCTTACCTGCCAGCAGTACGCTGCTGCTGCAATGACTGCTACTGCACT 241  
QY 370 GCGTTTCTCCTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 429  
DB 242 GCGTTTCTCCTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301  
QY 430 GAATAT 435  
DB 302 GGAAT 307

RESULT 4  
ABO99467  
ID ABO99467 standard; CDNA; 1158 BP.  
XX AC ABO99467;  
XX  
XX 25-FEB-2003 (first entry)  
XX Human coding sequence SEQ ID 200.  
XX Human; expressed sequence tag; EST; chromosome 6q16.1-q16.3;  
KW haematopoietic disorder; central nervous system disease; viral infection;  
KW peripheral nervous system disease; non-healing wound; infectious disease;

KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;  
KW fungal infection; autoimmune disorder; coagulation disorder; neotropic;  
KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;  
KW cytostatic; haemostatic; virucide; antibacterial; fungicide;  
KW immunostimulant; cerebroprotective; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200259260-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 16-NOV-2001; 2001WO-US42950.  
XX  
PR 17-NOV-2000; 2000US-0714936.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;  
PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-590824/63.  
DR N-PSDB; ABP64881.  
XX  
XX New isolated polynucleotide, useful in research, diagnostic or  
PT therapeutic methods, e.g. preventing or treating disorders involving  
PT aberrant protein expression or biological activity -  
XX  
PS Claim 1; SEQ ID 200; 394pp; English.  
XX

CC The present invention relates to novel human coding sequences  
CC (ABQ99268-ABQ99608) and proteins (ABP64682-ABP65022). The sequences are  
CC useful in therapeutic, diagnostic and research methods. The  
CC polynucleotides may be used in the field of molecular biology as  
CC hybridisation probes, primers for PCR, for chromosome and gene mapping,  
CC for the recombinant production of protein, or in generation of anti-sense  
CC DNA or RNA. The polynucleotides are useful in diagnostics as expressed  
CC sequence tags (ESTs) for identifying expressed genes or for physical  
CC mapping of the human genome. The proteins may be used as molecular weight  
CC markers, or as nutritional sources or supplements. The proteins may be  
CC used to maintain and expand cell population in a totipotent or  
CC pluripotent state useful for re-engineering damaged or diseased  
CC tissues, transplantation, manufacture of bio-pharmaceuticals or the  
CC development of bio-sensors. The polynucleotides and proteins are useful  
CC for preventing, treating or ameliorating disorders involving aberrant  
CC protein expression or biological activity, e.g. haematopoietic disorders,  
CC central/peripheral nervous system diseases, mechanical and traumatic  
CC disorders, non-healing wounds, immune deficiencies and disorders,  
CC infectious diseases caused by viral, bacterial or fungal infection,  
CC autoimmune disorders, allergic reactions and conditions, coagulation  
CC disorders, or cancer. The polynucleotide sequences of the invention were  
CC assembled from ESTs isolated mainly by sequencing by hybridisation, and  
CC in some cases, sequences obtained from one or more public databases.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1158 BP; 337 A; 238 C; 263 G; 320 T; 0 other;

Query Match 13.3%; Score 262.4; DB 24; Length 1158;  
Best Local Similarity 57.6%; Pred. No. 1.1e-65;  
Matches 470; Conservative 0; Mismatches 346; Indels 0; Gaps 0;  
QY 97 TTGACTGTGCAAAATGCAAGCGAGTCCCTGTATGACGCAATACATCCAGACACAGC 156  
DB 247 TTTTACTGTCATCTGTCAGCAGCATCTCTTGGAGAGAAATATGTACTAAGATGAC 306  
QY 157 GGCCCTACTGTGTGCCCTGCTATGACAAATACCTTTGCCAACACCTGTGCTGAGTCCAG 216  
DB 307 AGTCCATCTGTGTACATGTTATGATCGTGTATTTTCTTACTATTGCGAGGAATGCAA 366  
QY 217 CAGCTTATCGGGGATGACTCGAGGAGCTGTTCATGAGACCGCCCATTCACAGGCGC 276





DB 906 AGCGTTTACCGCTGTGGAGGACCAAGTATTACTGGGTGATGCTACAGAACTTTGCG 965  
QY 737 CACCTAAGTGACGAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGS---CAAGTATG 793  
DB 966 CCAAGAAGTGCTGGATGCAAGAACCCCATCACTGGGTTGGTAAAGGCTCCAGTGGG 1025  
QY 794 TGTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTGCGCCGCTGCTTAOCT 853  
DB 1026 TGGCCTATGAAGGACAATCTTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCGTGA 1085  
QY 854 CCCTGGTGGCCAGGGCTTGTACCGGATGAGACCAAGTCTGTGCGAGGCTTAGCC 913  
DB 1086 ATCTGGCCCAACAAGCGCTTGTGTTTCCACGAGGAGAGTGTATTGTCCGACTGTGCCA 1145  
QY 914 A 914  
DB 1146 A 1146

RESULT 6  
ID ABK34596/c  
XX ABK34596 standard; cDNA; 2340 BP.

AC ABK34596;  
DT 08-MAY-2002 (first entry)

DE Human cDNA for novel secreted protein, SEQ ID 365.

XX Human; ss; gene; secreted protein; immune deficiency; viral infection;  
KW bacterial infection; fungal infection; autoimmune disorder; burn;  
KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;  
KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;  
KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;  
KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;  
KW lymphoid cell deficiency.

XX Homo sapiens.

OS WO200177290-A2.

PN 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10295.

PF 06-APR-2000; 2000US-194941P.

PR (GENY) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;

PI Gulkota K, Graham JR;

XX WPI; 2002-179323/23.

XX Six hundred and twenty five polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders

PS Claim 1; Page 183-184; 339pp; English.

XX The invention relates to 625 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins, their complements and sequences that hybridise to them.  
CC Also included are a vector comprising the polynucleotide, a host cell  
CC transformed with the vector, the proteins encoded by the  
CC polynucleotides, antibodies that bind to the proteins and identification  
CC of modulators of the proteins or the expression of the polynucleotide.  
CC The polynucleotides can be used as probes for the identification  
CC and isolation of full length cDNA and genomic DNA. The polynucleotides  
CC and proteins can also be used as nutritional supplements. The protein

CC is useful in the treatment of various immune deficiencies and disorders  
CC such as viral infections, bacterial infections, fungal infections,  
CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,  
CC autoimmune thyroiditis and diabetes) and allergic reactions and  
CC conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment  
CC of burns, incisions and ulcers. The proteins are also useful for  
CC regulating haematopoiesis, for treating myeloid or lymphoid cell  
CC deficiencies. The present sequence is one of the 625 cDNA sequences  
CC encoding a secreted protein.

XX  
SQ Sequence 2340 BP; 621 A; 515 C; 554 G; 650 T; 0 other;

Query Match 13.0%; Score 255.4; DB 24; Length 2340;  
Best Local Similarity 57.3%; Pred. No. 1.7e-63;  
Matches 482; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

QY 77 TTGCCACCATGAGCGAGTCACTTTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGA 136  
DB 2210 TGGGCACCATGCGGAGAGTTTGAATGCCACTACTGAGGGATCCCTTGCAGGGAGA 2151  
QY 137 AGTACATCCAGACAGACAGCGCCCTACTGTGTGCTGCTATGACAAATACCTTTGCCA 196  
DB 2150 AGTATGTCAAAAGGATGCCACCACTGCTGCTGAAATGCTTTGACAAGTTCTGTGCCA 2091  
QY 197 ACACCTGTGTAGTGCCAGCAGCTATTCGGGCATGACTCGAGGGAGCTGTCTATGAAG 256  
DB 2090 ACACCTGTGTGAATGCGCGCAAGCCATCGGTGCGGACTCCCAAGAGGTGCACTATA 2031  
QY 257 ACCGCGATTTCACGAGGCTGCTTCGCTGCTGCTGCCAGCGCTCCTACCTAGCCGATG 316  
DB 2030 ACCGCTTCTGGCATGACACCTGCTTCGCTGCTGCAAGTGCCTTCCACCTTTGGCCA 1971  
QY 317 AACCTTCCACCTGCCAGGACAGTGTGCTGCTGCAATGACTGTCTACTGCACTGCTGTT 376  
DB 1970 AGACCTTTGTGCCAAGGACACAAAGATCCTGTGCAACAAGTGCACCACTCGGAGG 1911  
QY 377 CTTGCACTGCTCCCTGTTGGGAGACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
DB 1910 CCCCCAAGTGCAAGGGGTGCTTCAAGGCCATTTGTGCGAGGAGATCAAAACGTTGAG 1851  
QY 437 GAGGCGACATGCGCATGAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496  
DB 1850 AGGGACCGCTTGGCAAAAGTCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791  
QY 497 CCGCTTCTTTTGTGCCGACAAAGGCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTG 556  
DB 1790 CTGGAAGCTTCTTCCCTAAAGGGGAGGACTTCTACTGCTGCTGCTGCTGCTGCTGCTG 1731  
QY 557 TTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616  
DB 1730 TTGCCAAGCATTTGCTGGAAGTGCACAAAGGCGCATCATCTGTGAGGAATCACTTACC 1671  
QY 617 ATGAGCGCTGCGATCGAAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 676  
DB 1670 ATGAGCGCTGCGATCGAAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1611  
QY 677 AGCAGTTCCACCTCCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 736  
DB 1610 AGCGTTTCCACCGCTGTGTGAGGACCAAGTATTACTGCTGCTGCTGCTGCTGCTGCTGCTG 1551  
QY 737 CACCTAAGTGCAGCAGCTGCAAGCGCCCATCTAGGACTCGGTGGAGS---CAAGTATG 793  
DB 1550 CCAAGAAGTGTGCTGGATGCAAGAACCCCATCACTGCGGTTTGGTAAAGGCTCCAGTGG 1491  
QY 794 TGTCTTTGAAGACCGACACTGGCACCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853  
DB 1490 TGGCCTATGAAGGACAATCTTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCCGCTGA 1431  
QY 854 CCCTGGTGGCCAGGGCTTCTGTTACCGGATGGAGACCAAGTGTCTGTGCCAGGCTGTAGCC 913







Db 316 AGACCTTTGTGGCCAAAGGACAAGATCTGTGCAACAAGTGCACCACCTCGGAGGACT 375  
 Qy 377 CTTGCGAGTCTCCGCTTGTGGGAGACTGTCTATGCTGGTCCCGAAGCTGGATATG 436  
 Db 376 TCCCCAAGTGAAGGGGTGCTTCAAGGCCATTTGGCAGAGATCAAAAAGTGGAGTACA 435  
 Qy 437 GAGGCCAGACATGGCATGAGCACTGTCTCTGTGCAAGTGGTGTGAACAGCCACTGGGCT 496  
 Db 436 AGGGAGCCGCTGGCACAAGACTGTCTACCTGTAGTACTCAAGCAAGTCTATCGGA 495  
 Qy 497 CCGGTTCTTTGTGCCCGACAAGGGTGTCTACTGCGTGGCCCTGTATGAGAACAAAT 556  
 Db 496 CTGGAAGCTTCTTCCCTAAGGGGGAGGACTTCTACTGCGTGAATGCGCATGAGACCAAGT 555  
 Qy 557 TTGCTCTCTGCTGCGCGCTGAGCAAGACGCTGACACAGGCTGGAGTGACATACCGTG 616  
 Db 556 TGCCCAAGCATTCGGTGAAGTGACAAAGGCCATCTGAGGAGAACTACCTTACCAGG 615  
 Qy 617 ATCAGCCGTGGCATCGAGAAATGTCTGTGTGTATACCGGATGCCAGAGCCCTGGCAGGC 676  
 Db 616 ATCAGCCCTGGCATCGGATGCTTGTGTGTGTACTCTTAAAGAACTGGCTGGC 675  
 Qy 677 AGCAGTTACCTCCCGGATGAAGATCCCTACTGTGTGGCTGTCTTTGGAGAACTCTTTG 736  
 Db 676 AGCGTTTACCGCTGTGGAGGACCACTTACTGCGTGGATGCTACAAGAACTTCTGTG 735  
 Qy 737 CACCTAAGTGCACAGCTGCAAGCGCCCATCTGTAGGACTGCGTGGAGG---CAAGTATG 793  
 Db 736 CCAAGAAATGTGTGATGCAAGAACCCCATCTACTGGGTTTGTAAAGGATCCAGTGTG 795  
 Qy 794 TGTCTTTTGAAGACCGACACTGGCACCACAACTGTCTCTCTGCGCCCGCTGCTACCT 853  
 Db 796 TGGCTATGAAGGACATCTGCGCAGCACTACTGCTTCCACTGCACAAAATGCTCCGTGA 855  
 Qy 854 CCTGTGTGGCCAGGGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGGCTGTAGCC 913  
 Db 856 ATCTGGCCACAAGCGCTTTGTTTCCACCAGGAGCAAGTGTATTGTCCCGACTGTGCCA 915  
 Qy 914 A 914  
 Db 916 A 916

## RESULT 9

AAH23101

ID AAH23101 standard; DNA; 2310 BP.

XX AC AAH23101;

XX DT 17-SEP-2001 (first entry)

XX DE Osteoarthritis tissue-derived nucleic acid sequence #31.

XX KW Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;

XX KW wound healing; osteopathic; anti-arthritis; anti-inflammatory;

XX KW vulnarary; antibacterial; antiallergic; ds.

XX OS Homo sapiens.

XX XX WO200153531-A2.

XX XX 26-JUL-2001.

XX XX 18-JAN-2001; 2001WO-US000016.

XX XX 18-JAN-2000; 2000US-0176523.

XX XX (PHAA ) PHARMACIA CORP.

XX PA Phippard D, Vasanthakamur G, Dotson S, Ma X;

XX PI WPI; 2001-451914/48.

XX DR

XX

PT

PT

PT

XX

PS

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

SQ

Sequence 2310 BP; 632 A; 547 C; 506 G; 625 T; 0 other;

Query Match 12.4%; Score 244.4; DB 22; Length 2310;

Best Local Similarity 57.2%; Pred. No. 2.7e-60;

Matches 482; Conservative 0; Mismatches 356; Indels 4; Gaps 2;

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Qy

Db

Substantially purified protein, polypeptide or their fragments, used to identify a biologically active compound or composition and treat mammalian osteoarthritis

Claim 1; Page 116-118; 144pp; English.

Sequences AAH23071-23152 represent nucleic acid sequences derived from osteoarthritis tissues. The sequences are useful as probes and for the diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides and polypeptides of the invention are useful for generating diagnostic reagents, as targets for small molecule drug development, generation of therapeutics, and cloning genes. Specific antibodies are used to generate enzyme linked immunosorbent assays for detection of osteoarthritis. The invented molecules can be used to treat osteoarthritis or to analyse the disease-modifying activity of osteoarthritis drugs. Other disorders treatable using the nucleic acid sequences include atopic, inflammatory and infectious disorders e.g. Crohn's disease and sepsis, and wound healing.



DT 21-NOV-2001 (first entry)  
XX Human reproductive system related antigen cDNA SEQ ID NO: 1354.  
DE Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ss.  
KW Homo sapiens.  
XX WO200155320-A2.  
XX 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01339.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231422.  
PR 08-SEP-2000; 2000US-0231423.  
PR 08-SEP-2000; 2000US-0231424.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-02331968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI: 2001-465570/50.  
PI P-PSDB; AAM95383.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen -  
PT is used in preventing, treating or ameliorating a medical condition -  
XX  
PS Claim 1; SEQ ID NO 1354; 1297pp + Sequence Listing; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a coding sequence of the  
CC invention.  
XX  
SQ Sequence 830 BP; 234 A; 181 C; 197 G; 212 T; 6 other;

Query Match 11.8%; Score 232.6; DB 22; Length 830;  
Best Local Similarity 57.5%; Pred. No. 4.4e-57;  
Matches 409; Conservative 4; Mismatches 298; Indels 0; Gaps 0;

QY 202 TGTGCTGAGTGCACAGCTTATCGGGATGACTCGAGGAGCTGTTCTATGAAGACCGC 261  
DB 11  
4 TCGAGGAATGCAAAAACCAATTAATCTGATTAAGGATCTTTGTACAAAGACCGG 63  
QY 262 CATTTCCAGGAGGCTGCTCGCTGCTCCGCTGCCGCTCAGGCGTCACTACCGCATGACC 321  
DB 11  
64 CACTGGCATGAAGGATGCTTCAAGTGACCAATGCAATCACTCTTTGGTGAAAGGCT 123  
QY 322 TTCACCTGCCAGCAGTGAAGTCTGCTGCAATGACTGCTACTGAGTGCCTTTTCCTCG 381  
DB 11  
124 TTTGCTGCCAAGATGAGCGCTGCTGTCACGGAGTCTATCTAATGAGTCTCTCC 183  
QY 382 CAGTCTCGCTGTTGGGAGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441  
DB 11  
184 AAGTCTCTCCACTGCAAGGAGGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243  
QY 442 CAGACATGCGATGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501  
DB 11  
244 AACTACTGCGATGAWACCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 303  
QY 502 TCTTTTGTGCCGACAAAGGCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561  
DB 11  
304 CTTTGTGATCTCAAGAGAGTGGCAATATTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
QY 562 COTCCTGCGCCGCTGCGAGCAAGAGCTGTACAGAGGTTGGAGTACATACCGTATCAG 621  
DB 11  
364 CACTACTGCAACTTTTGTGAAGAAGGTGATACTTCAAGTGGGATAAATTTTGTGACCA 423  
QY 622 CGTGGCATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681  
DB 11  
424 CTATGGCATGAAGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
QY 682 TTCACCTCCGGGATGAAGATCCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741  
DB 11  
484 TTCATGTCGAGAGCACTATCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
QY 742 AAGTGCAGCAGTGCAGGCGCCCATCTAGTACTCGGTGGAGGCAAGTATGTGCTCTTT 801  
DB 11  
544 AAGTGTGAGCTGTTTCAACCCCATTAAGTGTCTCAGAGTGGCAAGTTATCTGCTTT 603  
QY 802 GAAGACCCAGACTGCGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861  
DB 11  
604 CAAGACAGCCAGTGGCATAGGGAATGCTTAACTGCGGGAAATGCTGCTGCTGCTGCTGCTG 663

QY 862 GCCCAGGGCTGCTACCGGATGGAGACCAAGTGCCTGCGAGGCTGTAGC 912  
DB 11  
664 GGTAAAGGCTCTCACCACCAAGAAATCTTCTGCCAAATGTGGC 714

RESULT 12  
AAS33146  
ID AAS33146 standard; CDNA; 830 BP.  
XX  
AC AAS33146;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE DNA encoding human secreted protein, Seq ID No 105.  
XX  
KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155326-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01347.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI: 2001-451931/48.  
DR P-PSDB; AAU20437.  
XX  
PT New nucleic acids and polypeptides, useful for diagnosing, preventing  
PT or treating medical conditions -  
XX  
PS Claim 1; SEQ ID No 105; 753pp; English.  
XX  
CC The invention relates to novel isolated nucleic acid molecules (I)  
CC encoding human secreted proteins (II). (I) and (II) are used to prevent,  
CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,  
CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in  
CC the prevention, treatment and diagnosis of diseases associated with  
CC inappropriate expression of secreted proteins. (I) and complementary  
CC sequences may also be used as DNA probes in diagnostic assays (e.g.  
CC polymerase chain reactions (PCR)) to detect and quantitate the presence  
CC of similar nucleic acid sequences in samples, and so which patients may  
CC be in need of restorative therapy. (II) may also be used as antigens in  
CC the production of antibodies and in assays to identify modulators  
CC (agonists and antagonists) of the expression and activity of the secreted  
CC proteins. The anti-(II) antibodies and antagonists may also be used to  
CC down regulate expression and activity of (II). The anti-(II) antibodies  
CC may also be used as diagnostic agents for detecting the presence of (II)  
CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The  
CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,  
CC angina and thrombosis), infections caused by bacteria, viruses and  
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),  
CC agonists, antagonists and antibodies can also be used to promote wound  
CC healing, maintain organs before transplantation, and support cell culture

CC of primary tissues. AAS33043-AAS33486 represent human secreted protein  
CC coding sequences, PCR primers, and related sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO  
CC at: ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 830 BP; 234 A; 181 C; 197 G; 212 T; 6 other;  
Query Match 11.8%; Score 232.6; DB 22; Length 830;  
Best Local Similarity 57.5%; Pred. No. 4.4e-57;  
Matches 409; Conservative 4; Mismatches 298; Indels 0; Gaps 0;  
QY 202 TGTGCTGAGTGCAGAGCTTATCGGGCATGATCGAGGAGCTGTCTATGAAGACCGC 261  
Db 4 TGGAGGAATGCAAAAACCAATGAATCTGATTCATGAAGATCTTGTTCACAAAGACCG 63  
QY 262 CATTTCCAGAGGGCTGCTCCCTGCTGCCGCTGCAGGCTCAGTCCGATGAACCC 321  
Db 64 CACTGGCATGAAGGATGCTTCAAGTGCACCAATGCAATCACTCTTTGTTGGAAAGCCT 123  
QY 322 TTCACCTGCCAGCAGTGCCTGCTGCAATGCTGCTACTGCTGAGTGGTTCCTCG 381  
Db 124 TTGTGTCNAAAGATGAGCCCTGCTGTCACGAGTGTCTTAAAGAGTCTCTCTCC 183  
QY 382 CAGTGTCCGCTTGTGGGAGACTGTCATGCTCCGCTGGTCCGGAAGCTGGAATATGGAGC 441  
Db 184 AAGTGTCTCCACTGCAAGAGGACCATCATGCTGCTTCCGCAAAATGGAATTAAGGA 243  
QY 442 CAGACATGTCAGTCACTGCTTCCTGTCAGTGGCTGTGAACAGCCTGCGCTCCCGT 501  
Db 244 AACTACTGGCATGAWACCTGTGTGTGTGAGAAATGGCGACAAGCTATAGSGACAAG 303  
QY 502 TCCTTTGTGCGCCAGCAGGCTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561  
Db 304 CTTTGTATCTCCAAAGAGTGCATATTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363  
QY 562 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621  
Db 364 CACTACTGCACTTTGTAAGAAGGTGATAAATTCAGTGGGATAACATTTTGTGACCAG 423  
QY 622 CCTGTCATGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681  
Db 424 CTATGCTAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483  
QY 682 TTCACCTCCGGGATGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741  
Db 484 TTCATGTCAGAGACGACTATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543  
QY 742 AAGTCAGAGCTGCAAGCCGCCATCGTAGGACTCGGTGGAGGCAAGTATGTCTCTTT 801  
Db 544 AAGTGTGTAGCTGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603  
QY 802 GAAGCCGACACTGGCACCACACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861  
Db 604 CAAGACGACGCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663  
QY 862 GGCCAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
Db 664 GGTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714

RESULT 13  
ABL96806  
ID ABL96806 standard; cdna; 830 BP.  
XX  
AC ABL96806;  
XX  
XX 21-JUN-2002 (first entry)  
XX Human testicular antigen encoding cdna SEQ ID NO: 474.  
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW

KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ss.  
XX Homo sapiens.  
XX WO200155317-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01329.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.







PT comprises several cDNAs that are differentially expressed in activated  
PT vascular tissue -  
XX  
PS Claim 1; Page -: 18pp; English.

XX This invention relates to a combination comprising several cDNAs that  
CC are differentially expressed in activated vascular tissue. The invention  
CC also discloses a high throughput method for detecting differentially  
CC expressed cDNAs in a sample. The cDNAs of the invention may have  
CC antiarteriosclerotic, cytostatic, cardiant; hypotensive; antidiabetic;  
CC synaenological; vasotropic and cerebroprotective activities and may be  
CC used in gene therapy. The cDNAs of the invention may be used in a  
CC high-throughput methods for detecting differential expression of one or  
CC more cDNAs in a sample, or screening several molecules or compounds to  
CC identify a molecule or compound that specifically binds a cDNA of the  
CC invention. A protein encoded by the cDNA may be used to screen several  
CC molecules or compounds to identify a ligand that specifically binds to  
CC the protein, or to produce or purify an antibody to the protein that can  
CC be used to detect a protein in a sample or purify a natural or  
CC recombinant protein from a sample. The nucleotides may be useful for  
CC diagnosing, staging, treating, or monitoring the progression of  
CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary  
CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-  
CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used  
CC for large-scale genetic or gene expression analysis of several new  
CC nucleic acid molecules. Antibodies to the proteins encoded by the  
CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic  
CC or acute diseases associated with abnormalities in the expression,  
CC amount or distribution of the protein. The present sequence  
CC represents a cDNA of the invention that is differentially expressed in  
CC activated vascular tissue.

CC Note: The sequence data for this patent did not form part of the  
CC specification, but was obtained in electronic format directly from USPTO  
CC at <http://seqdata.uspto.gov/sequence.html?docID=20020137081>.

XX  
SQ Sequence 2154 BP: 550 A; 553 C; 513 G; 536 T; 2 other;

Query Match 11.4%; Score 225; DB 25; Length 2154;  
Best Local Similarity 57.7%; Pred. No. 1.1e-54;  
Matches 402; Conservative 0; Mismatches 295; Indels 0; Gaps 0;

QY 77 TTGCCACCATGAGCGAGTCATTGTTGACTGTGCAAAATGCAACGAGTCCCTGTATGGACGCA 136  
DB 172 TGGGCACCATGCGGAGAAAGTTTGACTGTCCTACTGCGGGGATCCCTTGCAGGGGGAAGA 231  
QY 137 AGTACATCCAGACAGACAGCGCCCTACTGTGTCCTGTCTATGACAAATACCTTTTGCCA 196  
DB 232 AGTATGTCAAAGAGATGCCACCACTGCTGCTGAAATGCTTTGACAAGTTCTGTGCCA 291  
QY 197 ACACCTGTGCTGAGTGCCAGCAGCTTATCGGGCATGACTCGAGGGAGCTGTCTATGAAG 256  
DB 292 ACACCTGTGAGTATGCCGCAAGCCCATCGGTGCGGACTCCCAAGGAGGTGCACATATAAGA 351  
QY 257 ACCGCCATTTCACAGAGGCTGCTTCGCGTGTGCGCGTGCAGCGCTCCTAGCCGATG 316  
DB 352 ACCGCTTCTGGCATGACACCTGCTTCGCGTGTGCGCAAGTGCCTTCACCCCTTGGCCAATG 411  
QY 317 AACCTCTTCACCTGCGCAGGAGTGTGCTGCAATGACTGTCTACTGCTGCGGTGTTT 376  
DB 412 AGACCTTTGTGGCCAAAGGACAAAGATCCCTGTGTCACAAAGTGCCACCTCGGGAGGACT 471  
QY 377 COTCGAGTGTCCGCTTTGTGGGGAGACTGTGTCATGCTGGGTCCCGGAAGCTTGAATATG 436  
DB 472 CCCCCAAGTGCAAGGGGTGCTTCAAGGCCATTGTGGCAGGAGATCAAAACGTGGAGTACA 531  
QY 437 GAGGCCACATGCGCATGACGACTGCTTCCTGTGTCAGTGGCTGTGTAACAGCCACTGGGCT 496  
DB 532 AGGGAGCCGTGTGGCACAAGAGACTGCTTCACCTGTAGTAACTGCAAGCAAGTCAATCGGGA 591  
QY 497 CCGGTTCTTTGTGCGCAGCAAGGGTGTCTACTGCTGCGCTGCTATGAGAACAAAGT 556  
DB 592 CTGGAGAGCTTCTTCCCTAAAGGGGAGGACTTCTACTGCTGACTTGCCATGAGACCAAGT 651

QY 557 TTGCTCCTCGCTGCGCCGCTGCACGAAGACGCTGACACAGGGTGGAGTGACATACCCGTG 616  
DB 652 TTGCCAAGCATTTGCGTGAAGTGCAACAAGGCCATCACATCTGAGGAATCACTTACCAGG 711  
QY 617 ATCAGCCCTGGCATCGAAGATGCTGTGCTGTACCGGATGCCAGAGCGCCCTGGCAGGGC 676  
DB 712 ATCAGCCCTGGCATCGCGATTGCTTTGTGTGTACCTGTCTAAGAAGCTGGCTGGGC 771  
QY 677 AGCAGTTACCTCCCGGATGAAGATCCCTACTGTGTGCGCTGTTTGGAGAACTCTTTG 736  
DB 772 AGCGTTTACCCTGTGGAGGACCAGTATTAATCTGCGTGGATTGCTACAGAAGAACTTTG 831  
QY 737 CACCTAAGTGCAGCAGCTGCAGCGCCCATCGTAGG 773  
DB 832 CCAAGAAGTGTCTGGATGCAAGAACCCTCATCTGG 868

Search completed: August 27, 2003, 12:03:10  
Job time : 398 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 11:52:39 ; Search time 103 Seconds

(without alignments)  
8446.273 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 ggccgctctgcagtcggcag.....ctctgacgtgggggatgaa 1971

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	240	12.2	2254	4	US-09-016-434-1122
2	75.4	3.8	1488	4	US-09-124-238A-8
3	75.4	3.8	1488	4	US-09-721-975-8
4	75.4	3.8	1488	4	US-09-986-621-8
5	75.4	3.8	1597	1	US-08-166-316-1
6	75.4	3.8	1620	4	US-09-124-238A-32
7	75.4	3.8	1620	4	US-09-721-975-32
8	75.4	3.8	1620	4	US-09-986-621-32
9	75.4	3.8	1644	4	US-09-124-238A-9
10	75.4	3.8	1644	4	US-09-721-975-9
11	75.4	3.8	1644	4	US-09-986-621-9
12	75.4	3.8	1665	4	US-09-124-238A-33
13	75.4	3.8	1665	4	US-09-721-975-33
14	75.4	3.8	1665	4	US-09-986-621-33
15	75.4	3.8	1689	4	US-09-124-238A-22
16	75.4	3.8	1689	4	US-09-721-975-22
17	75.4	3.8	1689	4	US-09-986-621-22
18	68.4	3.5	232	4	US-09-016-434-386
19	66.8	3.4	1696	4	US-09-124-238A-2
20	66.8	3.4	1696	4	US-09-721-975-2
21	66.8	3.4	1696	4	US-09-986-621-2
22	66.6	3.4	2060	4	US-09-523-849-1
23	65.2	3.3	1674	1	US-08-889-402-7
24	65.2	3.3	1776	1	US-08-889-402-3
25	65.2	3.3	1776	1	US-08-889-402-5
26	65.2	3.3	1818	1	US-08-889-402-4
27	65.2	3.3	1818	1	US-08-889-402-6

28	56.8	2.9	7218	1	US-08-232-463-14	Sequence 14, Appl
29	53.8	2.7	1959	3	US-09-008-465-2	Sequence 2, Appl
30	53.8	2.7	1959	4	US-09-528-959-2	Sequence 2, Appl
31	51	2.6	2102	3	US-08-966-318-2	Sequence 2, Appl
32	51	2.6	2102	3	US-09-216-619-2	Sequence 2, Appl
33	50.2	2.5	231	4	US-09-439-313-455	Sequence 455, App
34	50.2	2.5	231	4	US-09-352-616A-455	Sequence 455, App
35	49.6	2.5	1866	4	US-09-205-258-66	Sequence 66, Appl
36	46.6	2.4	1166	4	US-09-072-596-323	Sequence 323, App
37	44.2	2.2	1234	4	US-09-166-350-24	Sequence 24, Appl
38	44.2	2.2	1246	1	US-08-289-441A-1	Sequence 1, Appl
39	44.2	2.2	1246	3	US-09-276-851-1	Sequence 1, Appl
40	43.6	2.2	915	4	US-09-740-235-10	Sequence 10, Appl
41	40.8	2.1	565	4	US-09-669-751-9	Sequence 9, Appl
42	40.8	2.1	749	4	US-09-669-751-26	Sequence 26, Appl
43	40.2	2.0	2955	4	US-09-252-991A-797	Sequence 797, App
44	37.8	1.9	694	3	US-09-328-111-631	Sequence 631, App
45	37.4	1.9	689	3	US-09-328-111-832	Sequence 832, App

ALIGNMENTS

RESULT 1

US-09-016-434-1122

; Sequence 1122, Application US/09016434

; Patent No. 6500938

; GENERAL INFORMATION:

; APPLICANT: Janice Au-Young

; APPLICANT: Jeffrey J. Sellhammer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

; TITLE OF INVENTION: PATHWAY GENE EXPRESSION

; NUMBER OF SEQUENCES: 1490

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/016,434

; FILING DATE: HEREWITH

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0002 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 1122:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2254 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GENBANK

; CLONE: g1381807

US-09-016-434-1122

Query Match 12.2%; Score 240; DB 4; Length 2254;  
Best Local Similarity 56.1%; Pred. No. 5.8e-59;

Matches	472;	Conservative	0;	Mismatches	366;	Indels	3;	Gaps	1;
Qy	77	TTGCCACCATGAGCGAGTCA	TTTTGACTGTGCAAAATCAACGAGTCCCTGTATGGACGCA	136					
Db	76	TGGSCACCATGGCGGAGAA	TTTGACTGCGCACTACTCGAGGATCCCTTCGACGGGAAGA	135					
Qy	137	AGTACATCCAGACAGACAGCGGCCCTTACTGTGTGCCCTGCTATGACAATACCTTTGCCA	196						
Db	136	AGTATGTGCAAAAGGATGGCCACACCTGCTGCCTGAAATGCTTTGACAAGTCTTGTGCCA	195						
Qy	197	ACACCTGTGCTGAGTGCAGCAGCTTATCGGGGATGACTCGAGGAGCTGTTCTATGAAG	256						
Db	196	ACACCTGTGTGGAATGCCGACGCCCATCGTGCGGACTCCAAGGAGTGCACATAAGA	255						
Qy	257	ACGCGCAATTCACGAGGGTGCTTCGCGTGTGCGCGCTGCCAGCGCTCACTAGCCGATG	316						
Db	256	ACCGCTTCTGGCATGACACCTGCTTCGCGTGTGCCAAGTGCCTTCAACCTTTGCCCAAT	315						
Qy	317	AACCCCTTCACTGCCAGGACAGTGAGTGTCTCTGCAATGACTGCTACTGCAGTGCCTTTT	376						
Db	316	AGACCTTTTGTGGCAAGGACACAAGAATCTCTGTGCACAGTGCACCACTNGGGAGGACT	375						
Qy	377	CTCTGCGAGTGTCCGCTTTGTGGGAGACGTCTCATGCTGGGTCCCGGAAGCTGGAATATG	436						
Db	376	TCGCCAAGTGCAGGGGTGCTTCAAGGCCATTTGTGGCAGGAGATCAAAAGTGGAGTACA	435						
Qy	437	GAGSCCAGACATGSCATGACCATGCTTCTCTGTGCACTGGCTGTGAACAGCCACTGGCT	496						
Db	436	AGGGACCGTCTGGCACAAAGACTGCTTTACCTGTAGTAGTACTCGAACCAAGTCAATCGGA	495						
Qy	497	CCCGTTCTTTTGTGCCGACAAAGGGTGCTCACTACTGCGTGCCCTGCTATGAGACAAAGT	556						
Db	496	CTGGAAGCTTCTTCCCTAAGGGGAGGACTTCTACTGCTGACTTGGCCATGAGACCAAGT	555						
Qy	557	TTGCTCTCGTGGCCCGCTGCAGCAAGACGCTGCACAGGTGGAGTGACATACCGTG	616						
Db	556	TGGCCAAGCATTGGTGAAGTGCACAAGGCCATCACTCTGAGGGAATCACTTACCAGG	615						
Qy	617	ATCAGCCGTGGCATTCGAGAATGTCTGTCTGTACCCGGATCCAGACGCCCTCGCAGGGC	676						
Db	616	ATCAGCCCTGGCATGCCGATTGCTTTGTGTGTGTACTCTGCTTAAGAAGCTGGCTGGGC	675						
Qy	677	AGCAGTTCACTCCCGGATCAAGATCCCTACTGTGTGGCCTGTTTGGAGAACTCTTTG	736						
Db	676	AGCTTTACCGCTGTGGAGACCAATTTACTGCGTGGATGCTACAAGAACTTCGTGG	735						
Qy	737	CACCTAAGTCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGG---CAAAGTATG	793						
Db	736	CCAAGAAGTGTGTGGATGCAAGAACCCTCACTGGTTTGTGTAAGGCTCCAGCTGTG	795						
Qy	794	TGTCCTTTGAAGACCGACACTGGACACAACTGCTTCTCTCGGCCCGCTGCTACCT	853						
Db	796	TGGCCTATGAAGGACAACTCTGCGACGACTACTGCTTCCACTGCAAAAAATGCTCCGTGA	855						
Qy	854	CCCTGTGGCGCAGGGCTTCGTACCGGATGGAGACCAAGTGCCTGCGCGGGGTGTAGCC	913						
Db	856	ATCTGGCCACACAGCGCTTTGTTTCCACAGGAGCAAGTGATTGTCCCGACTGTGCCA	915						
Qy	914	A	914						
Db	916	A	916						

? TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors  
 ?  
 ? FILE REFERENCE: 06148.0115  
 ?  
 ? CURRENT APPLICATION NUMBER: US/09/721,975  
 ?  
 ? CURRENT FILING DATE: 2000-11-27  
 ?  
 ? PRIOR APPLICATION NUMBER: US 09/124,238  
 ?  
 ? PRIOR FILING DATE: 1998-07-29  
 ?  
 ? PRIOR APPLICATION NUMBER: 60/054,219  
 ?  
 ? PRIOR FILING DATE: 1997-07-30  
 ?  
 ? PRIOR APPLICATION NUMBER: 60/080,407  
 ?  
 ? PRIOR FILING DATE: 1998-04-02







Db 1049 GCGAGATCATGACGGCCCTGAAGATGACCTGGCAGCTGCACCTTACCTGTGCTGCT 1108  
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTACCTCCCGGATGAAGATCCCTACTGTGG 715  
Db 1109 GCAAGAGCCCATCCGGAACAGGGCTTCTACATGGAGGAGGCGCTATTGGAGC 1168  
QY 716 CCGTGTGGAGAACTCTTTGGACCTTAAGTCAGCAGCTGCAAGGCCCCATCGTAGGAC 775  
Db 1169 GAGACTATGAGAAGATGTTGGCAGAAATGCCATGGCTGTGACTTCAAGATCG----- 1222  
QY 776 TCGGTGGAGCAAGTATGTCTTCTTGAAGCCGACACTGGCACCACACAACCTCTCTCT 835  
Db 1223 ACGCTGGGAGCGCTTCCCTGGAGGCCCTTCCAGCTGGCATGACACCTCTCTCTCT 1282  
QY 836 GGGCCCGCTGCTTCTACCTCCCTGGTGGCCAGGGCTTCTGATCCGGATGGAGCAAGTGC 895  
Db 1283 GTGCCATATGTCAGATCAACCTTGAAGGAAAGACCTTCTACTCCAAGAGGACAGGCCTC 1342  
QY 896 TCTGCCAGGC 906  
Db 1343 TCTGCAAGAGC 1353

## RESULT 10

US-09-721-975-9

; Sequence 9, Application US/09721975  
; Patent No. 6444803  
; GENERAL INFORMATION:  
; APPLICANT: Hair, Gregory A.  
; APPLICANT: Boden, Scott D.  
; TITLE OF INVENTION: No. 6444803el Bone Mineralization Proteins, DNA, Vectors,  
; FILE REFERENCE: Expression Systems  
; FILE REFERENCE: 06148.0115  
; CURRENT APPLICATION NUMBER: US/09/721,975  
; CURRENT FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: US 09/124,238  
; PRIOR FILING DATE: 1998-07-29  
; PRIOR APPLICATION NUMBER: 60/054,219  
; PRIOR FILING DATE: 1997-07-30  
; PRIOR APPLICATION NUMBER: 60/080,407  
; PRIOR FILING DATE: 1998-04-02  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: MS Word  
; SEQ ID NO 9  
; LENGTH: 1644  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-721-975-9

Query Match 3.8%; Score 75.4; DB 4; Length 1644;  
Best Local Similarity 48.7%; Pred. No. 9.8e-12;  
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;  
QY 416 GGTCCCGAAGCTGGAATATGGAGCCAGACATGGCATGAGCAGCTCTCTCTGTGCACTG 475  
Db 869 GGGCCCGTACTGTGTGGCTTGGGCCACGCGTACACCCGGAGGAGTTTGTGTGATGCC 928  
QY 476 GCTGTGAACAGCACTGGCTCCCGTCTTTTGTGCCGACAAAGGCTCTCACTACTGCG 535  
Db 929 AGTGTGGAGGTCTGGAAGGGTGGCTTCTTTGAGGAGAGAGGCGCATCTTCTGCTGCC 988  
QY 536 TGCCCTGTATGAGAACAAAGTTTGTCTCTCGCTGCGCCGCTGCAAGAGAGCGCTGACAC 595  
Db 989 CACCATGTATGACGTGGCTGTATGCACCCAGCTGTGCCAAGTGCAAGAAGATTACAG 1048  
QY 476 GCTGTGAACAGCACTGGCTCCCGTCTTTTGTGCCGACAAAGGCTCTCACTACTGCG 535  
Db 929 AGTGTGGAGGTCTGGAAGGGTGGCTTCTTTGAGGAGAGAGGCGCCATCTTCTGCC 988  
QY 536 TGCCCTGTATGAGAACAAAGTTTGTCTCTCGCTGCGCCGCTGCAAGAGAGCGCTGACAC 595  
Db 989 CACCATGTATGACGTGGCTGTATGCACCCAGCTGTGCCAAGTGCAAGAAGATTACAG 1048  
QY 596 AGGTGGAGTGCATACCGTGCATGATCGCTGATCGAGATGTCTGCTCTCTACCGGAT 655  
Db 1049 GCGAGATCATGACGCCCTTGAAGATGACCTGGCAGCTGCACCTTACCTGTGCTGCT 1108  
QY 656 GCCAGACGCCCTGGCAGGCGACAGTTTCACTCCCGGATGAAGATCCCTACTGTGG 715  
Db 1109 GCAAGACGCCCATCCGAACAGGGCTTCTACATGGAGGAGGCGCTATTCGCGAGC 1168  
QY 716 CCGTGTGGAGAACTCTTTGGACCTTAAGTCAGCAGCTGCAAGGCCCCATCGTAGGAC 775  
Db 1169 GAGACTATGAGAAGATGTTGGCAGAAATGCCATGGCTGTGACTTCAAGATCG----- 1222  
QY 776 TCGGTGGAGCAAGTATGTCTTCTTGAAGCCGACACTGGCACCACACAACCTCTCTCT 835  
Db 1223 ACGCTGGGAGCGCTTCTCTGGAGGCCCTTCTGAGGCGCTTCTGAGTGGCATGACCTGCTCT 1282



Qy	836	GCGCCGCTCCTTACTCCTCGTGGCGCAGGGCTTGTACCGATGGAGACCAAGTGC	895
Db	1283	GTCCGATATGTAGATCAACTGGAAGGAAGACCTTCTACTCCAAGAGGACAGGCCTC	1342
Qy	896	TCTGCCAGGGC	906
Db	1343	TCTGCAAGAGC	1353

**RESULT 12**

```

US-09-124-238A-33
; Sequence 33, Application US/09124238A
; Patent No. 6300127
; GENERAL INFORMATION:
; APPLICANT: Hair, Gregory A.
; APPLICANT: Boden, Scott D.
; TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,
; TITLE OF INVENTION: Expression Systems
; FILE REFERENCE: 06148.0115
; CURRENT APPLICATION NUMBER: US/09/124,238A
; CURRENT FILING DATE: 1998-07-29
; PRIOR APPLICATION NUMBER: 60/054,219
; PRIOR FILING DATE: 1997-07-30
; PRIOR APPLICATION NUMBER: 60/080,407
; PRIOR FILING DATE: 1998-04-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: MS Word
; SEQ ID NO 33
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-124-238A-33

```

## RESULT 14

US-09-986-621-33 ; Sequence 33, Application US/09986621  
; Patent No. 6521750  
; GENERAL INFORMATION:

APPLICANT: Hair, Gregory A.  
APPLICANT: Boden, Scott D.  
TITLE OF INVENTION: No. 6521750el Bone Mineralization Proteins, DNA, Vectors,  
FILE OF INVENTION: Expression Systems  
FILE REFERENCE: 06148.0115  
CURRENT APPLICATION NUMBER: US/09/986,621  
PRIOR FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: 09/124,238  
PRIOR FILING DATE: 1998-07-29  
PRIOR APPLICATION NUMBER: 60/080,407  
PRIOR FILING DATE: 1998-04-02  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: MS Word  
SEQ ID NO 33  
LENGTH: 1665  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-986-621-33

Query Match 3.8%; Score 75.4; DB 4; Length 1665;  
Best Local Similarity 48.7%; Pred. No. 9,8e-12;  
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;  
QY 416 GGTCCCGGAAGCTGGAATATGGAGCCAGACATGCGACATGACGACTGCTTCTGTGCGAGT 475  
DB 914 GGGGCGGCTACTGTGGCGCTTGGGCCACGGGTACACCCCGGAGGAGTTTGTGTGAGCC 973  
QY 476 GCTGTGAACAGCACTGGCTCCCGTCTTTTGTGCCGACAAAGGCTGCTCACTACTCGG 535  
DB 974 AGTGTGGGAAGTCTGGAAGAGGTGGCTTCTTTGAGGAGAGAGGCGCCATCTTCTGCC 1033  
QY 536 TGCCCTGCTATGAGAACAAAGTTTGTCTGCTGCGCGCGCTGCGACAGCGCTGACAC 595  
DB 1034 CACCATGCTATGACGTGCGCTATGACCCAGCTGTGCCAAGTGCAAGAGAGATTACAG 1093  
QY 596 AGGTGGAGTACATACCGTGTGATGACGCGTGGCATCGAGATGCTGCTGTACCGGAT 655  
DB 1094 GCGAGATCATGACGCCCTGAAGATGACCTGGCAGGTGCACTGCTTTACCTGTGCTGCT 1153  
QY 656 GCCAGACGCCCTGGCAGGGCAGCAGTTTCACTCCCGGGATGAAGATCCCTACTGTGTTG 715  
DB 1154 GCAAGACGCCCATCCGGAACAGGGCTTCTACATGGAGGAGGGCGTCCCTATTGCGAGC 1213  
QY 716 CTTGTTTGGAGAACTCTTTGCACTTAAGTGGCAGCAGCTGCAAGCGCCCGCCATGCTG 775  
DB 1214 GAGACTATGAGAAGATGTTTGGCAGAAATGCCATGGCTGTCAAGATCG----- 1267  
QY 776 TCGGTGGAGCAAGTATGCTGCTTTGAAGACCGACACTGGCACCACCACTGCTTCTCT 835  
DB 1268 ACGTGGGAGCGGCTTCTGAGGCGCTTCTGAGGCGCTTCTGAGGCGCTTCTGCTCT 1327  
QY 836 GCGCCCGCTGCTTCTACCTCCCTGTTGGCGCAGGGCTTCTGATCCCGGATGGAGACCAAGTGC 895  
DB 1328 GTGGATATGTCAGATCAACCTGGAAGGAAGACCTTCTACTCCAAAGGAGGACGCGCTC 1387  
QY 896 TCTGCCAGGC 906  
DB 1388 TCTGCAAGAGC 1398

RESULT 15  
US-09-124-238A-22  
Sequence 22, Application US/09124238A  
Patent No. 6300127  
GENERAL INFORMATION:  
APPLICANT: Hair, Gregory A.  
APPLICANT: Boden, Scott D.  
TITLE OF INVENTION: No. 6300127el Bone Mineralization Proteins, DNA, Vectors,  
FILE OF INVENTION: Expression Systems  
FILE REFERENCE: 06148.0115  
CURRENT APPLICATION NUMBER: US/09/124,238A  
CURRENT FILING DATE: 1998-07-29  
PRIOR APPLICATION NUMBER: 60/054,219

PRIOR FILING DATE: 1997-07-30  
PRIOR APPLICATION NUMBER: 60/080,407  
PRIOR FILING DATE: 1998-04-02  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: MS Word  
SEQ ID NO 22  
LENGTH: 1689  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-124-238A-22  
Query Match 3.8%; Score 75.4; DB 4; Length 1689;  
Best Local Similarity 48.7%; Pred. No. 9,9e-12;  
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;  
QY 416 GGTCCCGGAAGCTGGAATATGGAGCCAGACATGCGACATGACGACTGCTTCTGTGCGAGT 475  
DB 914 GGGGCGGCTACTGTGGCGCTTGGGCCACGGGTACACCCCGGAGGAGTTTGTGTGAGCC 973  
QY 476 GCTGTGAACAGCACTGGCTCCCGTCTTTTGTGCCGACAAAGGCTGCTCACTACTCGG 535  
DB 974 AGTGTGGGAAGTCTGGAAGAGGTGGCTTCTTTGAGGAGAGAGGCGCCATCTTCTGCC 1033  
QY 536 TGCCCTGCTATGAGAACAAAGTTTGTCTGCTGCGCGCGCTGCGACAGCGCTGACAC 595  
DB 1034 CACCATGCTATGACGTGCGCTATGACCCAGCTGTGCCAAGTGCAAGAGAGATTACAG 1093  
QY 596 AGGTGGAGTACATACCGTGTGATGACGCGTGGCATCGAGATGCTGCTGTACCGGAT 655  
DB 1094 GCGAGATCATGACGCCCTGAAGATGACCTGGCAGGTGCACTGCTTTACCTGTGCTGCT 1153  
QY 656 GCCAGACGCCCTGGCAGGGCAGCAGTTTCACTCCCGGGATGAAGATCCCTACTGTGTTG 715  
DB 1154 GCAAGACGCCCATCCGGAACAGGGCTTCTACATGGAGGAGGGCGTCCCTATTGCGAGC 1213  
QY 716 CTTGTTTGGAGAACTCTTTGCACTTAAGTGGCAGCAGCTGCAAGCGCCCGCCATGCTG 775  
DB 1214 GAGACTATGAGAAGATGTTTGGCAGAAATGCCATGGCTGTCAAGATCG----- 1267  
QY 776 TCGGTGGAGCAAGTATGCTGCTTTGAAGACCGACACTGGCACCACCACTGCTTCTCT 835  
DB 1268 ACGTGGGAGCGGCTTCTGAGGCGCTTCTGAGGCGCTTCTGAGGCGCTTCTGCTCT 1327  
QY 836 GCGCCCGCTGCTTCTACCTCCCTGTTGGCGCAGGGCTTCTGATCCCGGATGGAGACCAAGTGC 895  
DB 1328 GTGGATATGTCAGATCAACCTGGAAGGAAGACCTTCTACTCCAAAGGAGGACGCGCTC 1387  
QY 896 TCTGCCAGGC 906  
DB 1388 TCTGCAAGAGC 1398

Search completed: August 27, 2003, 14:13:45  
Job time : 106 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 13:24:15 ; Search time 375 Seconds  
(without alignments)

12058.582 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 ggccctgtgcagtcgagcag.....ctctgacgtggggatgaa 1971

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1533700 seqs, 1147125425 residues

\*Total number of hits satisfying chosen parameters: 3067400

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	429	21.8	431	11	US-09-918-995-36070
2	344.4	17.5	346	14	US-10-066-543-2286
3	252.2	12.8	2254	10	US-09-954-456-2244
4	244.4	12.4	2310	10	US-09-765-231A-31
5	232.6	11.8	830	11	US-09-764-891-1354
6	225	11.4	2154	13	US-10-044-090-260
7	152.8	7.8	580	10	US-09-833-381-786
8	117.6	6.0	445	11	US-09-918-995-4641
9	115.8	5.9	426	11	US-09-918-995-5309
10	112.4	5.7	424	11	US-09-918-995-4146
11	107.8	5.5	604	10	US-09-833-381-1725
12	99.8	5.1	724	10	US-09-833-381-1726
13	88.6	4.5	550	11	US-09-991-936-78
14	75.6	3.8	1486	9	US-09-854-864-3
15	75.4	3.8	1488	11	US-09-986-625-8
16	75.4	3.8	1620	11	US-09-986-625-32

#### ALIGNMENTS

##### RESULT 1

US-09-918-995-36070

; Sequence 36070, Application US/0918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 36070

; LENGTH: 431

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(431)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-36070

Query Match 21.8%; Score 429; DB 11; Length 431;  
Best Local Similarity 99.8%; Pred: No. 3.2e-124;  
Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	424	AAGCTGGAATGATGAGGCCAGACATGCGATGAGCAGCTGCTTCTGTGTCAGTGGCTGTGAA	483
Db	1	AAGCTGGAATGATGAGGCCAGACATGCGATGAGCAGCTGCTTCTGTGTCAGTGGCTGTGAA	60
Qy	484	CAGCACTGGCTCCCGTCTTTTGTGCCGACAGAGGTGCTACTACTGCTGCTCCCTGC	543
Db	61	CAGCACTGGCTCCCGTCTTTTGTGCCGACAGAGGTGCTACTACTGCTGCTCCCTGC	120
Qy	544	TATGAGAACAAAGTTTGTCTCGTGCCTGCGCCGCTGTCAGCAGAGCTGACACAGGCTGGA	603

Sequence 9, Appl  
Sequence 33, Appl  
Sequence 22, Appl  
Sequence 32, Appl  
Sequence 44, Appl  
Sequence 28, Appl  
Sequence 223, Appl  
Sequence 221, Appl  
Sequence 76, Appl  
Sequence 23603, A  
Sequence 37, Appl  
Sequence 7733, Ap  
Sequence 19419, A  
Sequence 8148, Ap  
Sequence 2701, Ap  
Sequence 3390, Ap  
Sequence 4613, Ap  
Sequence 2, Appl  
Sequence 1739, Ap  
Sequence 4879, Ap  
Sequence 4381, Ap  
Sequence 4829, Ap  
Sequence 4177, Ap  
Sequence 39, Appl  
Sequence 4885, Ap  
Sequence 4744, Ap  
Sequence 8634, Ap  
Sequence 11000, A



```
Db 376 TCCCAAGTGAAGGGGTGCTTCAAGGCCATTTGGCAGGAGATCAAAACGTGGAGTACA 435
Qy 437 GAGCCAGACATGACATGACACTGCTTCTCTGTCAGTGGTGTGAACAGCCACTGGGCT 496
Db 436 AGGGAGCGTCTGGCAAAAGACTGCTTCACTGTAGTAACTGCAACAAGTCAATCGGA 495
Qy 497 CCGTTCCTTTTGGCCCGCAAGAGGTGCTCACTACTGCGTCCCTGCTATGAGAACAAAT 556
Db 496 CTGGAAGCTTCTTCCCTTAAGGGGAGGACTTCTACTGCGTGAATGGCATGAGACCAAGT 555
Qy 557 TTGCTCTCTGCTGCGCCCGTGCAGCAAGAGCTGACACAGGGTGGAGTGACATACCGTG 616
Db 556 TGGCCAAAGCATTTGGTGAAGTGCAACAAAGGCCATCACATCTGGAGGAATCACTTTACCAG 615
Qy 617 ATCAGCCGTGGCATGAGAAATGCTGTGCTGTACCGGATGCGACAGCCCTGCGAGGC 676
Db 616 ATCAGCCCTGGCATGCGGATTTGCTTGTGTGTACTGCTTAAGAAGCTGCTGGC 675
Qy 677 AGCATTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTTTTGGAGAACTTTTG 736
Db 676 AGCGTTTACCGCTGTGGAGGACCAGTATTACTGCGTGGATTGCTACAAGAACTTCGTGG 735
Qy 737 CACCTAAGTGCAGAGCTGCAAGCGCCCATCTAGACTCGGTGGAGG---CAAGTATG 793
Db 736 CCAAGAAGTGTGTGGATGCAAGAACCCCATCACTGGGTTTGGTAAAGGCTCCAGTGTGG 795
Qy 794 TGTCTTTGAAGACCGACACTGGCACCACCAACTGCTTCTCTGCTGCGCCGCTCTACCT 853
Db 796 TGGCTATGAAGACATCTGGCAGCACTACTGCTTCCACTGCAAAAATGCTCCGTGA 855
Qy 854 CCGTGTGGCCAGGGCTCGTACCGGATGGAGACCAAGTGTCTGTGCGAGGGCTGTAGCC 913
Db 856 ATCTGGCCAACAAGCGCTTTGTTTCCACCAGGAGCAAGTGTATTGTCCCGACTGTGCCA 915
Qy 914 A 914
Db 916 A 916

RESULT 4
US-09-765-231A-31
; Sequence 31, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Phippard, Deborah
; APPLICANT: Vasanthakumari, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765,231A
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 31
; LENGTH: 2310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-31

Query Match 12.4%; Score 244.4; DB 10; Length 2310;
Best Local Similarity 57.2%; Pred. No. 5.7e-66;
Matches 482; Conservative 0; Mismatches 356; Indels 4; Gaps 2;

Qy 77 TTGCCACCATGAGCGAGTATTTGACTGTGCAAAATGCAACAGAGTCCCTGTATGGAGCA 136
Db 119 TGGCCACCATGAGCGAGTATTTGACTGTGCAAAATGCAACAGAGTCCCTGTATGGAGCA 178
Qy 137 AGTACATCCAGACAGAGCGGCCCTTACTGTGTGCGCTGATGACAAATACCTTTGCCA 196
Db 179 AGTATGTCAAAAGGATGGCCACCACTGCTGCCTGAAATGCTTTGACAAAGTCTGTGCCA 238
```

```
Qy 197 ACACCTGTGTGAGTGGCAGCAGCTTATCGGCGCATGACTCGAGGAGAGCTGTTCTATGAAG 256
Db 239 ACACCTGTGTGAATGCGCAAGCCATCGGTGGGACTTCCAGGAGGTGCACATATAAGA 298
Qy 257 ACCGCCATTTCCAGGAGGGTCTTCGCTGCTCGCTGCTCCAGCGCTCACT-AGCCGAT 315
Db 299 ACCGCTTCTGGCATGACACCTGCTTCGCTGTGCAAGTGTCTTACCCCTTGGGCCAAT 358
Qy 316 GAACCTTACCTCCAGGAGCAGTGAAGTGTCTGCAATGACTCTACTGTCAGTGCCTTT 375
Db 359 GAGACCTTTTGGCCAAAGGACAAAGATCTCTGCAACAAAGTGCACACCTCGGGAGGAC 418
Qy 376 TCCTGCGAGTCTCCGCTTGTGGGAGAGTGTCTATGCTTGGGTCCCGGAAGCTGGAATAT 435
Db 419 TCCCCCAAGTGCAGAGGGTGTCTTCAAGGCCATTTGGCAGGAGATCAAAAGTGGAGTAC 478
Qy 436 GGAGGCCAGACATGGCATGAGCACTGCTTCTGTGTCAGTGGCTGTGAACAGCCACTGGG 495
Db 479 AAGGGGACCGCTCTGGSCAAAGACTGCTTACCTGTGTAGTACTGCAAGCAAGTCACTGGG 538
Qy 496 TCCGCTTCTTTGTGCGCCGACAAAGGTGCTCACTACTGCGTGCCTTGTATGAGAAACAAG 555
Db 539 ACTGGAAGCTTCTTCCCTTAAAGGGAGGACTTCTACTGCGTGAATTCATGAGACCAAG 598
Qy 556 TTTGCTCTGCTGCTGCGCCGCTGCAGCAAGACGCTGACACAGGGTGGAGTGACATACCGT 615
Db 599 TTTGCCAAGCATTCGCTGAAGTGCACAAAGGCCATCACATCTGGAGGAATCACATTACCAG 658
Qy 616 GATCAGCCGTGGCATGAGAAATGCTGTGTGTACCGGATGCCAGAGCCAGCCCTGGCAGGG 675
Db 659 GATCAGCCGTGGCATGCGGATTTGCTTGTGTGTGTACTGCTCTAAAGAAAGTGGCTGGG 718
Qy 676 CAGCAGTTCACCTCCCGGATGAAGATCCCTACTGTGTGCGCTGTTTGGAGAACTCTTT 735
Db 719 CAGGTTTCCCGCTGTGGAGGACCAGTATTACTGCGTGGATTCATCAAGAACTTTGTG 778
Qy 736 GCACCTAAGTGCAGCAGCTGCAAGCGCCCATCTAGGACTCGGTGGAGG---CAAATAT 792
Db 779 GCCAAGAAGTGTGCTGGATGCAAGAACCCCATCACTGGTGTGTAAGGCTCCAGTGTG 838
Qy 793 GTGTCTTTTGAAGACCGACACTGGCACCACACACTGCTTCTCTGCGCCGCTGCTCTACC 852
Db 839 GTGGCTTATGAAGGACAAATCCTGGCAGGACTACTGCTTCCACTGCAAAAAATGCTCCGTG 898
Qy 853 TCCCTGTGGGCCAGGCTTGTGATCCGATGGAGCAAGAGTGTCTGTCAGGGCTGTAGC 912
Db 899 AATCTGGCCCAACAGCGCTTTGTTTCCACAGGAGCAAGTGTATTGTCCGACTGTGCC 958
Qy 913 CA 914
Db 959 AA 960

RESULT 5
US-09-764-891-1354
; Sequence 1354, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1354
; LENGTH: 830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-1354

Query Match 11.8%; Score 232.6; DB 11; Length 830;
```

Best Local Similarity 57.5%; Pred. No. 2e-62;  
Matches 409; Conservative 4; Mismatches 298; Indels 0; Gaps 0;

```
QY 202 TGTGCTGAGTCAGCAGCTTATCGGGCATGACTCGAGGAGCTGTCTCTATGAAGACCGC 261
  || || || || || || || || || || || || || || || || || || || || || ||
Db 4 TCGGAGGATGCAAAAACCAATGAATCTGATTTAGGATCTTTGTTACAAAGACCGG 63

QY 262 CATTTCCACGAGGCTGTTCGCTGCTGCCCTGCCAGCGCTCACTAGCCGATGAACCC 321
  || || || || || || || || || || || || || || || || || || || || || ||
Db 64 CACTGGCATGAAGGATGCTTCAAGTGCACCAAAATCAATCACTCTTTGGTGGAAAGCCT 123

QY 322 TTCACCTGCCAGCAGTGAAGTCTCTGCAATGACTGTACTGAGTCCGTTTTCCTCG 381
  || || || || || || || || || || || || || || || || || || || || || ||
Db 124 TTTGCTGCCAAGGATGAGCGCTGTGTGCACGAGTGTCTATTTCTAATGAGTGCCTCTC 183

QY 382 CAGTGTCTCCGCTTGTGGGAGACTGTCTATGCTCGGTCGCGGAAGCTGGAATATGGAGC 441
  || || || || || || || || || || || || || || || || || || || || || ||
Db 184 AAGTGTCTCACTGCAAGGACCAATCATGCTGTTCGCCGCAAAATGGAATTAAGGA 243

QY 442 CAGACATGSCATGAGCACTGTCTCTGTCAGTGTGTGAACAGCCACTGGGCTCCCGT 501
  || || || || || || || || || || || || || || || || || || || || || ||
Db 244 AACTACTGSCATGAWACCTGTGTTGTGTGAGAAATGCGGACAACCTATATGGACAAAG 303

QY 502 TCTTTTGTGCCGCAAGGCTGCTCACTACTGCTGCTGCTGCTGCTATGAGAACGTTTGTCT 561
  || || || || || || || || || || || || || || || || || || || || || ||
Db 304 CCTTTGATCTCCAAAGAGAGTGGCAATTAATTTGTGTGCCATGTTKTGAGAAGAGTTTGTCT 363

QY 562 CCTGCTGGGCGCTGCGAGCAGAGCTGTGACACAGGCTGAGTGCATACCTGATCAG 621
  || || || || || || || || || || || || || || || || || || || || || ||
Db 364 CACTACTGCACTTTGTGAAGGATGATCACTCAGTGGGATACATTTTGTGACCAAG 423

QY 622 CGTGGCTGAGAAATGTGTCTGTACCGGATGCCAGACGCCCTGCGAGGCGAGCAG 681
  || || || || || || || || || || || || || || || || || || || || || ||
Db 424 CTATGGCATAAAGAGTGTCTTCTGTGTAGTGTCTGAGAAAGATCTCTGTGAAGAACAG 483

QY 682 TTCACCTCCCGGATGAAGATCCCTACTGTGTGGCTGTTTGTGGAGAACTCTTTGCACT 741
  || || || || || || || || || || || || || || || || || || || || || ||
Db 484 TTCATGTCCAGAGAGCACTATCCATCTCTGCTGAGCTGTGTACAAACCACTTTATGCCAAC 543

QY 742 AAGTGCAGCAGCTGCAAGCGCCCATCGTAGGACTCGGTGGAGGAAGTATGTCTCTTT 801
  || || || || || || || || || || || || || || || || || || || || || ||
Db 544 AAGTGTGAGCTGTCTCAAAACCCATAGTGTCTCAGAGGTGCCAAGTTTATCTGCTTT 603

QY 802 GAAGACCGACACTGGCACCACAACTGCTTCTCTGCGCCGCTGCTCTACCTCCCTGCTG 861
  || || || || || || || || || || || || || || || || || || || || || ||
Db 604 CAAGACAGCACTGGCATAGCAATGCTTTAACTGCGGGAATGCTCTGCTCTCTGTTGGT 663

QY 862 GGCACGGCTTCTGACCGGATGAGACCAAGTGTCTCTGCGAGGCTGTAGC 912
  || || || || || || || || || || || || || || || || || || || || || ||
Db 664 GGTAAAGGCTTCTGACCCAGAACCAAGGAATCTTCTGCCAAAAATGTGGC 714
```

## RESULT 6

US-10-044-090-260  
; Sequence 260, Application US/10044090  
; Publication No. US20020137081A1  
; GENERAL INFORMATION:  
; APPLICANT: Olga Bandman  
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION  
; FILE REFERENCE: PA-0028 US  
; CURRENT APPLICATION NUMBER: US/10/044,090  
; CURRENT FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 850  
; SOFTWARE: PERL Program  
; SEQ ID NO 260  
; LENGTH: 2154  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20020137081A1 361583.16  
; NAME/KEY: unsure  
; LOCATION: 2106, 2150

; OTHER INFORMATION: a, t, c, g, or other  
US-10-044-090-260

```
Query Match 11.4%; Score 225; DB 13; Length 2154;  
Best Local Similarity 57.7%; Pred. No. 6.8e-60;  
Matches 402; Conservative 0; Mismatches 295; Indels 0; Gaps 0;  
QY 77 TTGCCACCATGAGCGAGTCAATTTGACTGTGCAAAAATGCAACGAGTCCCTGTATGACGCA 136  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 172 TGGGCACCATGCGGAGAGTTTGTACTGCTACTGTGCAAGGATCCCTTGCAGGGAAGA 231  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 137 AGTACATCCAGACACAGACGCGCCCTACTGTGTGCCCCTGCTATGACAAATCACTTTGGCA 196  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 232 AGTATGTGCAAAAGATGCGCCACCACTGCTGTGCTGAAATGCTTTGACAAAGTTCTGTGCA 291  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 197 ACACCTGTGCTCAGTGCAGCAGCTTATCGGCATGACTCGAGGAGCTGTTCTTCTATGAAG 256  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 292 ACACCTGTGTGAATGCGCAAGCCCATCGGTGGGACTCCCAAGAGGTGCACTATPAAGA 351  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 257 ACCGCAATTTCCACGAGGCTGCTTCCGCTGTGCGCTGCCAGGCTCACTAGCCGATG 316  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 352 ACCGCTTCTGGCATGACACCTGCTTCCGCTGTGCAAGTGCCTTCAACCCCTTGGCCAAATG 411  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 317 AACCTTCCACCTGCCAGGACAGTGAAGTCTGCTGCAATGACTGTCTACTGCACTGCGTTT 376  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 412 AGACCTTTTGTGCCAAGGACAAACAAGATCCTGTGCAACAAGTGCACCACTCGGGAGGACT 471  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 377 CCTCAGTGTCTCCCTTGTGGGAGACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 472 CCCCAGTGTGAAGGGTGTCTTCAAGGCCATTTGGCAGAGATCAAAACGTGGAGTACA 531  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 437 GAGGCGACAGATGSCATGAGCACTGCTTCTCTGCTGAGTGGCTGTGAACAGCCACTGGGCT 496  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 532 AGGGGACCGCTGCGCACAAAGACTGCTTCAACCTGTAGTAACCTGCAAGCAAGTCATCGGGA 591  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 497 CCGCTTCTTTGTGCGCCGACAAAGGTGCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 592 CTGGAAGCTTCTTCTCCCTAAAGGGGAGGACTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 651  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 557 TTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 616  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 652 TTGCCAAGCATTTGCTGGAAGTGCACAAAGCCCATCATCTCTGGAGGAATCACTTACCAGG 711  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 617 ATCAGCCGTGGCATCGAAGATGCTGTGCTGTACCGGATGCCAGAGCGCCCTGGCAGGCG 676  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 712 ATCAGCCCTGGCATGCCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 771  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 677 AGCAGTTCACTTCCCGGATGAAGATCCCTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 772 AGCGTTTCAACCGCTGTGGAGGACCAAGTATTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 831  
  || || || || || || || || || || || || || || || || || || || || || ||  
QY 737 CACCTAAGTGCAGCAGCTGCAAGCGCCCTCATCTAGG 773  
  || || || || || || || || || || || || || || || || || || || || || ||  
Db 832 CCAAGAAGTGTGCTGGATGCAAGAAGCCCACTGAG 868
```

## RESULT 7

US-09-833-381-786/c  
; Sequence 786, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 786  
; LENGTH: 580  
; TYPE: DNA









Db 511 AAACCAATTGAATGATCTAAGGATCTTTGTACAAAGACCGGCACTGGCATGAAGGA 570  
Qy 277 TGTTCGGCTGCGCGCTCCGAGCTACTAGCCGATCAACCCCTTCACTGCCAGGAC 336  
Db 571 TGTTCNAAGTGACCAAAATCAATCACTCTTTGGTGAAGAAGCTTTTGTGCGCAAGGAT 630  
Qy 337 AGTGAGCTGCTCTGCAATGACTGCTACTGTCAGTGCCTTTTCCCTCGCAGTCTCGCCTGTGT 396  
Db 631 GAGCGCTGCTGTGCGAGGAGTCTATTCTAACGAGTGTCTTNCRAAGTCTTNCACGTGN 690  
Qy 397 GGGGAGCTCTCATGCTGCTGGTCCCGGAAGCTGGA 431  
Db 691 AAGANGAC-CTCATGCTGCTGCTCCCGCAAAATGGA 724

## RESULT 13

US-09-991-936-78

; Sequence 78, Application US/09991936

; Publication No. US20030073827A1

; GENERAL INFORMATION:

; APPLICANT: Brandt, Kevin S.

; APPLICANT: Gaines, Patrick J.

; APPLICANT: Stinchcomb, Dan T.

; APPLICANT: Wisniewski, Nancy

; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE

; FILE REFERENCE: FC-6-C1

; CURRENT APPLICATION NUMBER: US/09/991,936

; CURRENT FILING DATE: 2001-11-21

; PRIOR APPLICATION NUMBER: US/09/543,668

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: 60/128,704

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 1959

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 78

; LENGTH: 550

; TYPE: DNA

; ORGANISM: Ctenocephalides felis

US-09-991-936-78

Query Match 4.58; Score 88.6; DB 11; Length 550;  
Best Local Similarity 52.08; Pred. No. 2.9e-17;  
Matches 199; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

Qy 97 TTTGACTGTCAAAATGCAAGAGTCCCTGTATGGAGCGCAAGTACATCCAGACAGACG 156  
Db 71 TTCTGCTGTGGCAATCGGAGAGTCACTGACCGGCCAACCGCTACGTACTCCGCGAGAA 130  
Qy 157 GGCCCTTACTGTGTGCCCTGTATGACAATACCTTTGCCAACACCTGTGTGTGAGTGCCAG 216  
Db 131 CATCTTACTGATCAAGTGTCTAGGAGCGCTCTTTCCAAACACCTCGGAGGATGCGAGC 190  
Qy 217 CAGCTTATCGGGATGACTCGAGGAGCTGTTTATGAAGACCGCCATTTCCAGAGGGC 276  
Db 191 AAGATCATTTGGCATTTGATTTCAAGGACTTATCTTACAAAGAAAGCATTTGCGATGAGGCA 250  
Qy 277 TGCTTCGGCTGCTGCCCTGCCAGCGTCACTAGCCGATGAACCTTCACTGCCAGGAC 336  
Db 251 TGTTCCTGTGTGAGTAATGCGCGTATCTCTCGTCGATAAACAGATTGCGGAAGTAATTG 310  
Qy 337 AGTGAGCTGCTCTGCATGACTGCTACTGAGTGCCTTTTCCCGAGTGTCTCGCTGTGT 396  
Db 311 GACAAATCTACTGTGGAATGCTGTATGATGCCCAATTCGTTCCAGGTGTGATGGTGC 370  
Qy 397 GGGGAGACTGCTATGCTGCTGGTCCCGGAAGCTGGAATATGGAGCCAGACATGCGATGAG 456  
Db 371 GCGGAATCTTCGTCGGGTACTAAAAAATGGAGTACAAAACACTGCTCAATGGCATGAA 430  
Qy 457 CACTGCTCTCTGTCAGTGGCTG 479  
Db 431 AAGTGTTCTGCTGTGTGTG 453

## RESULT 14

US-09-854-864-3

; Sequence 3, Application US/09854864

; Patent No. US20020081296A1

; GENERAL INFORMATION:

; APPLICANT: THEILL, LARS EYDE

; APPLICANT: YU, GANG

; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCMA,

; FILE REFERENCE: BLIS/AGP-3, AND TACI

; CURRENT APPLICATION NUMBER: US/09/854,864

; CURRENT FILING DATE: 2001-09-11

; PRIOR APPLICATION NUMBER: US 60/204,039

; PRIOR FILING DATE: 2000-05-12

; PRIOR APPLICATION NUMBER: US 60/214,591

; PRIOR FILING DATE: 2000-06-27

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 1486

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-854-864-3

Query Match 3.88; Score 75.6; DB 9; Length 1486;  
Best Local Similarity 56.28; Pred. No. 5.1e-13;  
Matches 163; Conservative 0; Mismatches 124; Indels 3; Gaps 1;

Qy 628 CATCGAGAATGCTGTGGTCTGTACCGGATGCCAGACGCCCCCTGGCAGGSCAGCAGTTTCACC 687  
Db 1 CATGCCGAGTGTCTTGTGTGTACCTGCTCTAAGAAGCTGGCTGGSCAGCGTTTCACC 60  
Qy 688 TCCCGGATGAGATCCCTACTGTGTGCGCTCTTTGGAGAACTCTTTGCACCTAAGTGC 747  
Db 61 GCTGTGGAGGACCACTATTACTGCTGCTGATTTGTGACAGAACTTTGTGGCCCAAGATGT 120  
Qy 748 AGCAGCTGCAAGCGCCCATCTAGTACCTGCTAGGAGTCCGTTGGAGG---CAAGTATGTCTCTTTGAA 804  
Db 121 GCTGTGCAAGAACCACCATCTACTGGGTTGGTAAGGCTCCAGTGTGGCTATGAA 180  
Qy 805 GACCGACACTGGCACCACAACTGCTTCTCTCCGCGCTGCTCTACCTCCCTCGTGGGC 864  
Db 181 GGCAATCTTGGCAGCACTACTGCTTCCACTGCACAAAAAATGCTCCGTGAATCTGGCCAA 240  
Qy 865 CAGGCTTCGTACCGGATGGAGACCAAGTCTCTGCCAGGCTGTAGCCA 914  
Db 241 AAGCGCTTGTATTTCATAATGACAGGTGTATTGCGCTGACTGTGCCAA 290

## RESULT 15

US-09-986-625-8

; Sequence 8, Application US/09986625

; Publication No. US20030125248A1

; GENERAL INFORMATION:

; APPLICANT: Hair, Gregory A.

; APPLICANT: Boden, Scott D.

; TITLE OF INVENTION: No. US20030125248A1el Bone Mineralization Proteins, DNA, Vecto

; FILE REFERENCE: 06148.0115

; CURRENT APPLICATION NUMBER: US/09/986,625

; PRIOR APPLICATION NUMBER: US 09/124,238

; PRIOR FILING DATE: 1998-07-29

; PRIOR APPLICATION NUMBER: 60/054,219

; PRIOR FILING DATE: 1997-07-30

; PRIOR APPLICATION NUMBER: 60/080,407

; PRIOR FILING DATE: 1998-04-02

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: MS Word

; SEQ ID NO 8

; LENGTH: 1488

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-986-625-8

Query Match      3.8%; Score 75.4; DB 11; Length 1488;
Best Local Similarity 48.7%; Pred. No. 5.8e-13;
Matches 239; Conservative 0; Mismatches 246; Indels 6; Gaps 1;

QY 416 GGTCCCGGAAGCTGGAATATGGAGCCGACATGCGACATGCTTCTGTGTCAGTG 475
   || || || || || || || || || || || || || || || || || || ||
Db 713 GGGGCGCTACTGTGTGGGCTGGCCACGCGTACACCCGGAGGAGTTGTGTAGCC 772
   || || || || || || || || || || || || || || || || || || ||

QY 476 GCTGTGAACAGCACTGGGCTCCCGTTCTTTGTGCCGACAAAGGTGCTACTACTGCG 535
   || || || || || || || || || || || || || || || || || || ||
Db 773 AGTGTGGGAAGTCTCTGGAAGAGGTGGCTTCTTTGAGGAGAAGGGCGCATCTTCTGCC 832
   || || || || || || || || || || || || || || || || || || ||

QY 536 TCCCTGTCTATGAGAACAAAGTTTGTCTCTCGCTGGCGCGCTGCGAAGACGCTGACAC 595
   || || || || || || || || || || || || || || || || || || ||
Db 833 CACCATGTATGACGTGCGCTATGACCCGAGCTGTGCCAAGTGCAAGAAGATTACAG 892
   || || || || || || || || || || || || || || || || || || ||

QY 596 AGGGTGGAGTACATACCGTGATCAGCGTGGCATCGAGATGCTCTGTGTACCGGAT 655
   || || || || || || || || || || || || || || || || || || ||
Db 893 GCGAGATCATGACGCCCTGAAGATGACCTGGCAGCTGCACCTGTACCTGTGTGCT 952
   || || || || || || || || || || || || || || || || || || ||

QY 656 GCCAGACGCCCTGGCAGGGCAGAGTTTCACTCCCGGGATGAAGATCCCTACTGTGTGG 715
   || || || || || || || || || || || || || || || || || || ||
Db 953 GCAAGACGCCCATCCGGAACAGGGCTTCTACATGGAGGAGGGCGTGCCTATTGCGAGC 1012
   || || || || || || || || || || || || || || || || || || ||

QY 716 CCTGTTTGGAGAACTCTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCGTAGGAC 775
   || || || || || || || || || || || || || || || || || || ||
Db 1013 GAGACTATGAGAAGATGTTGGCAGAAATGCCATGGCTGTGACTTCAAGATCG----- 1066
   || || || || || || || || || || || || || || || || || || ||

QY 776 TCGGTGGAGGCAAGTATGTCTCTTTGAAGACCGACACTGGCACCACAACTGCTTCTCT 835
   || || || || || || || || || || || || || || || || || || ||
Db 1067 AGGCTGGGACCGCTTCTCTGGAGGCCCTGGGCTTCAGCTGGCATGACACCTGCTTCGTCT 1126
   || || || || || || || || || || || || || || || || || || ||

QY 836 GCGCCCGTGTCTTACCTCCCTGGTGGCCAGGGCTTGTACCCGATGGAGACCAAGTGC 895
   || || || || || || || || || || || || || || || || || || ||
Db 1127 GTGCGATATGTGAGATCAACCTGGAAGGAAGACCTTCTACTCCAAGGAAGACAGGCCCTC 1186
   || || || || || || || || || || || || || || || || || || ||

QY 896 TCTGCCAGGGC 906
   || || || || || ||
Db 1187 TCTCAAGAGC 1197
   || || || || || ||
```

Search completed: August 27, 2003, 15:32:24  
Job time : 376 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 27, 2003, 11:50:19 ; Search time 2837 Seconds

(without alignments)

16885.486 Million cell updates/sec

Title: US-09-786-135-2

Perfect score: 1971

Sequence: 1 gdtccgtctgcagtcgcag.....ctcttgacgtggggatgaa 1971

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estnu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_man.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_pbg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	988.2	50.1	1201	13	BX385643
2	970.8	49.3	1172	9	AL527609
3	957	48.6	1201	9	AL523732
4	951.6	48.3	1201	13	BX437971

C	5	951.4	48.3	1082	9	AL526160
C	6	942.6	47.8	1201	9	AL528549
C	7	939.4	47.7	1201	9	AL525660
C	8	922.8	46.8	1200	13	BX345083
C	9	917.6	46.6	1201	9	AL526865
C	10	903	45.8	1201	9	AL523733
C	11	900.6	45.7	1200	9	AL526734
C	12	894.6	45.4	1201	13	BX385642
C	13	894.2	45.4	1116	9	AL551558
C	14	894.2	45.4	1201	13	BX445770
C	15	888.4	45.1	1034	12	BM480273
C	16	888.4	45.1	1044	12	BM557043
C	17	878.4	44.6	1201	9	AL563163
C	18	865.4	43.9	1096	9	AL550299
C	19	861.4	43.7	979	12	BM545780
C	20	836.8	42.5	1201	13	BX324699
C	21	830	42.1	877	9	AL521288
C	22	826.6	41.9	997	12	BI757781
C	23	816.4	41.4	1201	9	AL528415
C	24	812.8	41.2	1201	9	AL528520
C	25	811	41.1	933	9	AL521265
C	26	810.8	41.1	883	12	BM008985
C	27	810.6	41.1	1201	13	BX445771
C	28	807.2	41.0	853	12	BI822670
C	29	796.4	40.4	1201	9	AL528550
C	30	795.8	40.4	1071	9	AL550271
C	31	768.4	39.0	1043	9	AL525946
C	32	767.8	39.0	1125	9	AL562940
C	33	767.2	38.9	773	14	CA447438
C	34	765	38.8	1191	13	BX355333
C	35	764.2	38.8	855	12	BM010854
C	36	760.6	38.6	909	13	BQ675737
C	37	760.6	38.6	955	12	BG826380
C	38	756.8	38.4	1201	9	AL528387
C	39	756.2	38.4	1044	12	BM924990
C	40	747.6	37.9	1165	9	AL526900
C	41	746	37.8	865	12	BI908150
C	42	745	37.8	909	13	BX374168
C	43	744.6	37.8	1201	9	AL526613
C	44	741.8	37.6	891	13	BX372660
C	45	739.8	37.5	1145	9	AL551594

## ALIGNMENTS

RESULT 1

BX385643

LOCUS

DEFINITION

CDNA clone CS0DC014YA02 5-PRIME, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

-ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX385643 1201 bp mRNA linear EST 08-MAY-2003  
BX385643 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC014YA02 5-PRIME, mRNA sequence.

BX385643

BX385643.1 GI:30440553

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSIAC004ZC02QPI&cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSLAC004C02QPL.

## FEATURES

source

```
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC014YA02"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/clone="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      232 a   369 c   323 g   257 t   20 others
ORIGIN
Query Match      50.1%; Score 988.2; DB 13; Length 1201;
Best Local Similarity 96.5%; Pred. No. 1e-240;
Matches 1041; Conservative 9; Mismatches 24; Indels 5; Gaps 4;
QY 1 GGTCCGTCTGCAGTCGGCAGCTCGCCGGCAGCTCGCTCGCCCGCCGCTTGCGCCGGGCC 60
DB 68 GTTCGCTGTGAGTCGGCAGCTCGCCGGCAGCTCGCTCGCCCGCCGCTTGCGCCGGGCC 127
QY 61 GCAGGGTTCTTCGCCCTTGCACCACATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAG 120
DB 128 GCAGTNTNCTCTCCCTTGGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAG 187
QY 121 TCCGTGATGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGTCCTGCTAT 180
DB 188 TCCGTGATGACGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGTCCTGCTAT 247
QY 181 GACATACCTTGGCAACACCTGTGCTGAGTGCCAGCAGCTTATCGGCATGACTCGAGG 240
DB 248 GACATACCTTGGCAACACCTGTGCTGAGTGCCAGCAGCTTATCGGCATGACTCGAGG 307
QY 241 GAGCTGTTCTATGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTGCGCTGCCAG 300
DB 308 GAGCTGTTCTATGAAGACCGCATTTCCACGAGGCTGCTTCCGCTGCTGCGCTGCCAG 367
QY 301 CGCTCAGTACGATGACCTTACCTGCCAGGACACTGAGCTGCTGCAAGTACATGCTGC 360
DB 368 CGCTCAGTACGATGACCTTACCTGCCAGGACACTGAGCTGCTGCAAGTACATGCTGC 427
QY 361 TACTGACGTGGTTTCTCCTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 428 TACTGACGTGGTTTCTCCTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487
QY 421 CGGAAGCTGGAATATGAGGCGCAGACATGCGATGAGCAGTCTGCTGCTGCTGCTGCTG 480
DB 488 CGGAAGCTGGAATATGAGGCGCAGACATGCGATGAGCAGTCTGCTGCTGCTGCTGCTG 547
QY 481 GAACAGCCACT - GGCTCCCGTCTTTTGTGCGCCAGAGGTGCTCACTACTGCTGCTGC 539
DB 548 GAACAGCCACTGGGCTCCCGTCTTTTGTGCGCCAGAGGTGCTCACTACTGCTGCTGC 607
QY 540 CTGCTATGAGAACAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
DB 608 CTGCTATGAGAACAGTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667
QY 600 TGGAGTGACATACCTGATGAGCGTGGCATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 659
DB 668 TGGAGTGACATACCTGATGAGCGTGGCATGAGATGCTGCTGCTGCTGCTGCTGCTGCT 727
QY 660 GACGCCCTGCGAGGCGCAGCTTACCTCCCGGATGAAGTCCCTACTGTGTGGCCTG 719
DB 728 GACGCCCTGCGAGGCGCAGCTTACCTCCCGGATGAAGTCCCTACTGTGTGGCCTG 787
QY 720 TTTTGGAGAACTCTTTGACCACTAAGTGACGAGCTGCAAGCGCCGCCCTACTGAGGACTCG 779
DB 788 TTTTGGAGAACTCTTTGACCACTAAGTGACGAGCTGCAAGCGCCGCCCTACTGAGGACTCG 847
QY 780 TGGAGGCAAGTATGTGCTCTTTGAAGACCGGACACTGGCACCACCAACTGCTCTCTCT 839
```

```
DB 848 TGGAGGCAAGTATGTGCTCTTTGAAGACCGCAGACTGGCACCACAACTGCTTCTCCTGCC 907
QY 840 CGCTGCTCTACCTCCCTGGTGGCCAGGCGCTTCTACCGATGAGACC - AAGTGCCTCT 898
DB 908 CGCTGCTCTACCTCCCTGGTGGCCAGGCGCTTCTACCGATGAGACC - AAGTGCCTCT 967
QY 899 GCAGGGCTGTAGCCAGGCGCTTAAGCCAGGCGCTTCTACCGATGAGACC - AAGTGCCTCT 958
DB 968 GCCAGGGCTGTAG - CAGCAGGGGCTTAAGCCAGGCGCTTCTACCGATGAGACC - AAGTGCCTCT 1026
QY 959 CCACGGGCGCCAGGACTGTGGCTCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCG 1018
DB 1027 MCACGGGCGCCAGGACTGTGGCTCTTTTCTAAACACCTCTTTGGGACTMACCTCCCGCC 1086
QY 1019 CAAAAAATGGGTCTCTTCTGGGCTCCAGGATGTCTCCCACTCCAGCATCCCCAA 1077
DB 1087 CAAAAAATGGGTCTCTTCTGGGCTCAGGATGTCTT - CCACATCAGCATCCCCAA 1143

RESULT 2
LOCUS      AL527609      1172 bp      mRNA      linear      EST 23-MAY-2003
DEFINITION      AL527609 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION      AL527609
VERSION        AL527609.2  GI:31065460
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 1172)
AUTHORS      Li W.B., Gruber C., Jessee J. and Pollayes D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL
COMMENT      Unpublished
On Feb 13, 2001 this sequence version replaced gi:12791102.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 888.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC024CC120P1&cluster=888.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC024CC120P1.
FEATURES
Location/Qualifiers
1..1172
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YF23"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/clone="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      224 a   364 c   317 g   252 t   15 others
ORIGIN
Query Match      49.3%; Score 970.8; DB 9; Length 1172;
Best Local Similarity 96.9%; Pred. No. 2.8e-236;
Matches 1029; Conservative 4; Mismatches 21; Indels 8; Gaps 4;
QY 1 GGTCCGTCTGCAGTCGGCAGCTCGCCGGCAGCTCGCTCGCCCGCCGCTTGCGCCGGGCC 60
DB 70 GGTCCGTCTGCAGTCGGCAGCTCGCCGGCAGCTCGCTCGCCCGCCGCTTGCGCCGGGCC 129
QY 61 GCAGGGTTCTCTCCCTTGGCCACCATGAGCGAGTCAATTGACTGTGCAAAATGCAACGAG 120
```



```

Db 623 CTAAGCCAGGGCTCTGGACCCAGGCTTTCCATACACAGGGCCAGGACTGTGGCTCT 564
QY 984 TTTCTAAACCACTCTGGAGTCAAGCTCCCGCCGCCAAAATAATGGGTCTCTCTGGG 1043
Db 563 TTTCTAAACCACTCTGGAGTCAAGCTCCCGCCGCCAAAATAATGGGTCTCTCTGGG 504
QY 1044 CTCAGGATGTTCTCCCACTCCAGCATCCCAAACTGGTACTCCCTGACCCAGGCCGCC 1103
Db 503 CTCAGGATGTTCTCCCACTCCAGCATCCCAAACTGGTACTCCCTGACCCAGGCCGCC 444
QY 1104 RATCTTGGCTTTACAGAGCTCCATGAGTCAAGCCGCCCTCCCAACAGCTGGACTCAG 1163
Db 443 AATCTGG--TCTTACAGAGCTCCATGAGTCAAGCCGCCCTCCBYTACACCTGTACTCGG 386
QY 1164 AATTCACCCCTCCCTCGAGTCTGGGTTCCAGACTGAGTCTCTCCCAAACTCAGGSC 1223
Db 385 KATTCACCCCTCCCTCGAGTCTGGGTTCCCACTGAGTCTCTCCCAAACTCAGGSC 326
QY 1224 TCTAGACCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCTTTAAATCTAGACTT 1283
Db 325 TCTAGACCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCTTTAAATCTAGACTT 266
QY 1284 CTCCTTATAGTTTCAGGTCTCTATGGTGCCTGGAGTCCCTTGAAGTGGAGTCTTC 1343
Db 265 CTCCTTATAGTTTCAGGTCTCTATGGTGCCTGGAGTCCCTTGAAGTGGAGTCTTC 206
QY 1344 TCAGGCTTGACGTGCCCCACCCCAATCCCGGGGTTGAGGCTGTGGGGCAGCAGATCAGG 1403
Db 205 TCAGGCTTGACGTGCCCCACCCCAATCCCGGGTGGAGCTGTGGGGCAGCAGATCAGG 146
QY 1404 AGCCACATGATAAGGGGCCCTAGGTACAGGTGTGCTGCCAGCAGGTGCCACCCAGTGT 1463
Db 145 AGCCACATGATAAGGGGCCCTAGGTACAGGTGTGCTGCCAGCAGGTGCCACCCAGTGT 86
QY 1464 CTTCTCATTTTATTTTCACTCATTTTCCCATAGATGGCAGAGGGTGAGATTGGCTC 1523
Db 85 CTTCTCATTTTATTTTCACTCATTTTCCCATAGATGGCAGAGGGTGAGATTGGCTC 26
QY 1524 ATCCCCCTCCAGATTCGCAATAA 1548
Db 25 ATCCCCCTTCMAGATTCGCAATAA 1

```

## RESULT 4

## BX437971

## LOCUS

DEFINITION BX437971 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone

CS0DA002YG10 5-PRIME, mRNA sequence.

ACCESSION BX437971

VERSION BX437971.1 GI:30785652

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1201)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DA002BD050pl&cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DA002BD050Pl.

Location/Qualifiers

1..1201

## FEATURES

## source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA002YG10"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 236 a 373 c 315 g 257 t 20 others
ORIGIN

```

```

Query Match 48 3%; Score 951.6; DB 13; Length 1201;
Best Local Similarity 97.7%; Pred. No. 2.2e-231;
Matches 993; Conservative 3; Mismatches 14; Indels 6; Gaps 3;
QY 22 CTCGCCGAGCTCGCTCGGCCCGCCGCTTGGCCCGCCGCGCTCTCTCTCCCTTGGC 81
Db 64 CTCGCCGAGCTCGCTCGGCCCGCCGCTTGGCCCGCCGCGCTCTCTCTCCCTTGGC 122
QY 82 ACCATGAGGAGTCAATTTGACTGTGCAAAATGCAAGAGTCCCTGTATGGACGCAATAC 141
Db 123 ACCATGAGGAGTCAATTTGACTGTGCAAAATGCAAGAGTCCCTGTATGGACGCAATAC 182
QY 142 ATCCAGACAGACAGCGGCCCTTACTGTGTGCTGCTGCTATGACAATACCTTTGCCAACCC 201
Db 183 ATCCAGACAGACAGCGGCCCTTACTGTGTGCTGCTGCTATGACAATACCTTTGCCAACCC 242
QY 202 TGTGCTGAGTGCACAGACTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAGACCGC 261
Db 243 TGTGCTGAGTGCACAGACTTATCGGGCATGACTCGAGGAGCTGTTCTATGAAGACCGC 302
QY 262 CATTTCCAGAGGGCTGCTTCGGCTGCTGCCCTGCCAGCGCTCACTAGCCGATGACCC 321
Db 303 CATTTCCAGAGGGCTGCTTCGGCTGCTGCCCTGCCAGCGCTCACTAGCCGATGACCC 362
QY 322 TTCACCTGCCAGGACAGTGAAGTGTCTGCAATGACTGCTACTGCAAGTGTCTTCTCTCG 381
Db 363 TTCACCTGCCAGGACAGTGAAGTGTCTGCAATGACTGCTACTGCAAGTGTCTTCTCTCG 422
QY 382 CAGTCTCCGCTTGTGGGAGACTGCTCATGCTGCTGGTCCCGAAGCTGGAATATGAGGCG 441
Db 423 CAGTCTCCGCTTGTGGGAGACTGCTCATGCTGCTGGTCCCGAAGCTGGAATATGAGGCG 482
QY 442 CAGACATGSCATGAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
Db 483 CAGACATGSCATGAGCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
QY 502 TCTTTTGTGCCCGACAAGGGTGTCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
Db 543 TCTTTTGTGCCCGACAAGGGTGTCTCACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
QY 562 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
Db 603 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
QY 622 CCGTGGCATCGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 681
Db 663 CCGTGGCATCGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
QY 682 TTCACCTCCCGGATGAAGATCCCTTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 723 TTCACCTCCCGGATGAAGATCCCTTACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
QY 742 AAGTGCAGAGTGCAGCGGCCCATCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
Db 783 AAGTGCAGAGTGCAGCGGCCCATCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 838
QY 802 GAAGACCCACACTGCGACCAACACTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
Db 839 GAAGACCCACACTGCGACCAACACTGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898

```





JOURNAL  
COMMENT

Unpublished  
On Feb 13, 2001 this sequence version replaced gi:12792042.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 888.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DD001BC02NP1&cluster=888.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DD001BC02NP1.

## FEATURES

source

1. 1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DD001YB04"  
/tissue\_type="NEUROBLASTOMA COT 50-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 253 a 301 c 362 g 249 t 36 others

Query Match 47.8%; Score 942.6; DB 9; Length 1201;  
Best Local Similarity 96.4%; Pred. No. 4.3e-229;  
Matches 1008; Conservative 16; Mismatches 16; Indels 6; Gaps 6;

```

QY 475 GCCTGTGACACGACCTGGCTCCCGTCTTTGTGCGCGACAAAGGGTCTCACTACTGC 534
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1071 KGGTGTGACAGCCACTGGSTYCCCG-TCTTTTGTGCGCGAC-ARGGTGCTYAACTMTGC 1014
QY 535 GTGCCCTGCTATGAGAACAAAGTTTCTCTCTGCTGCGCGCGTGTGACGAAAGCGTGTGACA 594
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1013 GTGCCCTGCTATGAGAACAAAGTAACWCCTY-CGCGCGCGCTGCGACGAAGCGTGTGACA 955
QY 595 CAGGTGTGAGTACATACCGTGATCAGCGTGGCATCAGATGTCTGGTGTGTACCGGA 654
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
954 CAGGTGTGAGTACATACCGTGATCAGCGTGGCATCAGATGTCTGGTGTGTACCGGA 895
QY 655 TGCCAGACCCCTGCGCAGGCGAGCAGTTCACCTCCCGGATGAAGATCCCTACTGTGTG 714
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
894 TGCCAGACCCCTGCGCAGGCGAGCAGTTCACCTCCCGGATGAAGATCCCTACTGTGTG 835
QY 715 GCCTGTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCTCATCGTAGGA 774
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
834 GCCTGTTTGGAGAACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCTCATCGTAGGA 775
QY 775 CTCGGTGTGAGCAAGTATGTCTCTTTGAAGACCGACACTGGCACCACCAACTGCTTCC 834
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
774 CTCGGTGTGAGCAAGTATGTCTCTTTGAAGACCGACACTGGCACCACCAACTGCTTCC 715
QY 835 TCGCCCGCTGCTGTACCTCTCTGTGTGGCCAGGCTTCGTACCGGATGGAGACCAAGTG 894
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
714 TCGCCCGCTGCTGTACCTCTCTGTGTGGCCAGGCTTCGTACCGGATGGAGACCAAGTG 655
QY 895 CTCGTGCCAGGGCTGTGAGCCAGGCGCCCTTAAGCCAGGGTCTCTGGACCCAGGCTTTC 954
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
654 CTCGTGCCAGGGCTGTGAGCCAGGCGCCCTTAAGCCAGGGTCTCTGGACCCAGGCTTTC 595
QY 955 CATACCAGGGCCAGGACTGTGGTCTCTTTCTAAACCACTCTGGGACTCAGTCCCC 1014
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
594 CATACCAGGGCCAGGACTGTGGTCTCTTTCTAAACCACTCTGGGACTCAGTCCCC 535
QY 1015 CGCCAAAAAATGGGTCTCTCTCTGGGCTCCAGATGTCTCCCGCACTCCAGCATCCC 1074
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
534 CGCCAAAAAATGGGTCTCTCTCTGGGCTCCAGATGTCTCTCCCGCACTCCAGCATCCC 475

```

```

QY 1075 CAAATGGTACTCCTGTGACCCAGGCCCCCAATCCTCTGGGCTCTTACAGAGCCCTCCATGAGT 1134
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
474 CAAATGGTACTCCTGTGACCCAGGCCCCCAATCCT-GGCTCTTACAGAGCCCTCCATGAGT 416
QY 1135 CAAGCCCCCTCCACACACTGACCTGCAGCTCCAGAAATCACCTCTCCCTGCGAGTCTGGTTC 1194
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
415 CAAGCCCCCTCCACACAMMRAACTCCAGAAATCACCTCTCCCTGCGAGTCTGGTTC 356
QY 1195 CAGACTGAGTCTCTCTCCCAATCAGGGCTGTAGACCCGAGCCCTCCAAACCTGCACTCT 1254
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
355 CAGACTGAGTCTCTCTCCCAATCAGGGCTGTAGACCCGAGCCCTCCAAACCTGCACTCT 296
QY 1255 GGGACTTGGCCCCCTTAAATCTAGACTTCTCTTATAGTTTCAGGTTCTCTATGGGTG 1314
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
295 GGGACTTGGCCCCCTTAAATCTAGACTTCTCTTATAGTTTCAGGTTCTCTATGGGTG 236
QY 1315 CTTGGAGTCTCTTGAAGTGGACTGTCTCAGGCTTGACCTGCGCCACCCCATCCCGC 1374
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 CTTGGAGTCTCTTGAAGTGGACTGTCTCAGGCTTGACCTGCGCCACCCCATCCCGC 176
QY 1375 GGTTCAGGCTGTGGGGGCGAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGTAC-AG 1433
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
175 GGTTCAGGCTGTGGGGGCGAGCAGATCAGGAGCCCACTGATAAGGGGCCCTAGGTACAG 116
QY 1434 GGTCTGCCAGCAGCTGCCACCGAGTGT-CTTCTCAFTTTATTTCAGCTCCATTTGC 1492
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
115 GGTCTGCCAGCAGCTGCCACCGAGTGTACATCTCATTTATTCNCTCCATTTGC 56
QY 1493 CCATAGATGGGCGAGAGGGGTGAGATT 1518
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 CCATAGATGGGCGAGAGGGGTGAGATT 30

```

## RESULT 7

AL525660/c

LOCUS

DEFINITION

AL525660 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

ACCESSION

AL525660

VERSION

AL525660.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1201)

AUTHORS

Li W.B., Gruber C., Jessee J. and Polayes D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished

COMMENT

On Feb 13, 2001 this sequence version replaced gi:12789153.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 888.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DD0012CA04NP1&amp;cluster=888.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD0012CA04NP1.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DD001YB07"

/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V







QY 1309 TGGTGCCTGGGAAGTCCTTTAAAGTGGAGCTGTCTCAGCTTGAGCTGCCGCCACCCCAT 1368  
 Db 232 TGGTGCCTGGGAAGTCCTTTAAAGTGGAGCTGTCTCAGCTTGAGCTGCCGCCACCCCAT 173  
 QY 1369 CCCCGCGTTGAGCGTGGGGGACAGATCAGGAGCCACCTGATGAAGGGGCCCTAGGG 1428  
 Db 172 CCCCGCGTTGAGCGTGGGGGACAGATCAGGAGCCACCTGATGAAGGGGCCCTAGGG 113  
 QY 1429 TACAGGGTGTGCCAGCAGGTGGCCACCGAGTGTCTTCTCATTTTATTTACGCTCCATT 1488  
 Db 112 TACAGGGTGTGCCAGCAGGTGGCCACCGAGTGTCTTCTCATTTTATTTACGCTCCATT 53  
 QY 1489 TTGCCCATAGATGGGAGAGGGGTGAGATTGGCTCATGCCCTTCACGATTTCT 1541  
 Db 52 TTGCCCATAGATGGGAGAGGGGTGA---TTGCTCATGCCCTTBNNYCTCT 3

## RESULT 10

AL523733

LOCUS

DEFINITION

AL523733 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DC003YC05 5-PRIME, mRNA sequence.

ACCESSION

AL523733

VERSION

AL523733.2

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1201 bp mRNA linear EST 22-MAY-2003  
 AL523733 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
 cDNA clone CS0DC003YC05 5-PRIME, mRNA sequence.  
 AL523733  
 AL523733.2 GI:31041994  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 On Feb 13, 2001 this sequence version replaced gi:12787226.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 888.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DC003AB03QP1&cluster=888.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DC003AB03QP1.  
 Location/Qualifiers  
 1. .1201

## FEATURES

source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC003YC05"  
 /tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
 /notes="First strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 239 a 364 c 311 g 248 t 39 others

## BASE COUNT

ORIGIN

Query Match 45.8%; Score 903; DB 9; Length 1201;

Best Local Similarity 92.2%; Pred. No. 5.3e-219;

Matches 989; Conservative 24; Mismatches 49; Indels 11; Gaps 6;

QY 1 GGTCCCTGTCAGTCGGCAGCTCGCCGAGCTCGCTCGGCCCGCGCTTGGCCCGGCC 60

Db 70 GTTCCTCTGACGTCGGCAGCTCGCCGAGCTCGCTCGGCCCGCGCTTGGCCCGGCC 129

QY 61 GCAGGGTTCTCTCCCTTGGCCAGCATGAGGAGTCAATTGACTGTGCAAAATGCAACGAG 120

Db 130 GCA--GTTCTCTCCCTTGGCCAGCATGAGGAGTCAATTGACTGTGCAAAATGCAACGAG 187

QY 121 TCCTGTATGGAGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGCTAT 180  
 Db 188 TCCTGTATGGAGCAAGTACATCCAGACAGACAGCGGCCCTACTGTGTGCCCTGCTAT 247  
 QY 181 GACAATACCTTTGCCAACACCTGTGTAGTGCAGCAGCTTATCGGGCATGACTCGAGG 240  
 Db 248 GACAATACCTTTGCCAACACCTGTGTAGTGCAGCAGCTTATCGGGCATGACTCGAGG 307  
 QY 241 GAGCTGTTCTATGAAGACCGCCATTTCCACAGGGGTGCTTCGCTGCTCCGCTGCCAG 300  
 Db 308 GAGCTGTTCTATGAAGACCGCCATTTCCACAGGGGTGCTTCGCTGCTCCGCTGCCAG 367  
 QY 301 CGCTCACTAGCCGATGAACCTTCACTGCGCAGACAGTGAAGTCTCTCAATGACTGC 360  
 Db 368 CGCTCACTAGCCGATGAACCTTCACTGCGCAGACAGTGAAGTCTCTCAATGACTGC 427  
 QY 361 TACTGAGTGCCTTTTCTCGAGTGTCTCGCTTGTGGGAGAGTGTCTCATGCTGGTCC 420  
 Db 428 WACTGAGTGCCTTTTCTCGAGTGTCTCGCTTGTGGGAGAGTGTCTCATGCTGGTCC 487  
 QY 421 CGGAAGCTGGAATATGAGGCCAGACATGGCATGAGCAGTGCCTTCTGTGAGTGGCTGT 480  
 Db 488 CGGAAGCTGGAATATGAGGCCAGACATGGCATGAGCAGTGCCTTCTGTGAGTGGCTGT 547  
 QY 481 GAACAGCCACTGGGCTCCCGTCTTTGTGCGCGACAAGGGTGTCTCACTACTGCGTGCCC 540  
 Db 548 GAACAGCCA-TGGGCTCCCGTCTTTGTGCGCGACAAGGGTGTCTCACTACTGCGTGCCC 606  
 QY 541 TGTATGAGAACAAAGTTTGTCTCGTGGCGCCGCTGCAGCAGAGCTGCACACAGGTT 600  
 Db 607 TGTATGAGAACAAAGTTTGTCTCGTGGCGCCGCTGCAGCAGAGCTGCACACAGGTT 666  
 QY 601 GGAGTGACATACCGTATGATGAGCGTGGCATGAGAAATGTCTGTGTACCGGATGCCAG 660  
 Db 667 GGAGTGACATACCGTATGATGAGCGTGGCATGAGAAATGTCTGTGTACCGGATGCCAG 726  
 QY 661 ACSCCCTGCGAGGCGAGCAGTTTCACTCCCGGATGAAGATCCCTACTGTGTGGCCTGT 720  
 Db 727 AMGCCCTGCGAGGCGAGCAGTTTCACTCCCGGATGAAGATCCCTACTGTGTGGCCTGT 786  
 QY 721 TTGGAGAACTCTTTGCACCTAAGTGCAGAGCTGCAAGCGCCCATCTGTGAGGACTCGGT 780  
 Db 787 TTGGAGAACTCTTTGCACCTAAGTGCAGAGCTGCAAGCGCCCATCTGTGAGGACTCGGT 846  
 QY 781 GGAGGCAAGTATGTCTTGAAGACCGACATGGCACCACCACTGCTTCTCTCGCGCC 840  
 Db 847 GGAGGCAAGTATGTCTTGAAGACCGACATGGCACCACCACTGCTTCTCTCGCGCC 906  
 QY 841 CGCTGCTCTACCTCCCTGTTGGGCGCAGGCTTGTACCGGATGGAGACCAAGTCTCTGC 900  
 Db 907 CGCTGCTCTACCTCCCTGTTGGGCGCAGGCTTGTACCGGATGGAGACCAAGTCTCTGC 966  
 QY 901 CAGGGGTGTAGCCAGCAGGCGCCCTAAGCCAGGGCTCTTGGACCCAGGCTTTCCCAATCC 960  
 Db 967 CAGGGCTKTAG-CAGGCAGGCGCCCTAAG-CAGGGCTCTTGGACCCAGGCTTTCC--ATAC 1022  
 QY 961 ACGGGCCCCAGGACTGTGGCTCTTTCTTAACACCACTCTGGGACTGAGTCCCGGCCCA 1020  
 Db 1023 MACGGCCAGGACTGTGGCTCTTTTAAACACC-----TGTGGCGTTCCTTCTCTCTCTCT 1078  
 QY 1021 AAAAAAATGGGTCTCTCTTGGGCTCCAGGATTTGTCTCCCACTCCAGCATCC 1073  
 Db 1079 AATKGTYCTTCTGSSCCAGRATTTTTCOMCTCASCAYCCMAACTGKMCYCY 1131

## RESULT 11

AL526734/c

LOCUS

DEFINITION

AL526734 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DC020YH03 3-PRIME, mRNA sequence.

AL526734

ACCESSION

AL526734

VERSION

AL526734.2

GI:31064592

EST.



primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT	266 a	286 c	371 g	218 t	60 others
ORIGIN					
Query Match	45.4%; Score 894.6; DB 13; Length 1201;				
Best Local Similarity	92.7%; Pred. No. 7.3e-217;				
Matches 1002;	Conservative 37; Mismatches 31; Indels 11; Gaps 10;				
Qy	437	GAGCCAGACATGGCATGACGACGCTGCTCTGTCGACATGGCTGTGAACAGCCACTGGGCT	496		
Db	1096	GRGSCARASMTTGGATTASACTTSTCTCTGTSAGTGG--TGTGACGCCACTGGGCT	1039		
Qy	497	CCGCTCTCTTTGTCGCCGACAAAGGGTCTCACTACTCGCTGCTGATGAGAACAGT	556		
Db	1038	CCSKVTTTTTGTCCTCCMACAACGGGTGCTCASCCTGTCGACCTGATGGAARAKT	979		
Qy	557	TTGCTCTCTGCTGCGCCGCTGACGACGAGCGCT--GACACAGGTGGAGTGACATACCGT	615		
Db	978	TTGCTCTCT--TCGCGCCGCTGTCAGCAAGAGCGSTKGACACAGGKGAGTGACATACCGT	920		
Qy	616	-GATCAGCCCTGGCATCG--AGAATGCTGTGCTGTACTC--GGATGTCAGACGCCCT--TGCC	671		
Db	919	GGATCAGCCCTGGCATCGRGAATGCTGTGCTAGACCGGGATGCCAKACGCCCTCGGC	860		
Qy	672	AGGCGACAGCTTCACCTCCCGGATGAAGATCCCTACTGCTGTCGCTGTTTGGAGAA-C	730		
Db	859	AGGCGACAGCTTCACCTCCCGGATGAAGATCCCTACTGCTGTCGCTGTTTGGAGAAC	800		
Qy	731	TCCTTTCACCTTAAGTGCAGCAGCTGCAAGCGCCCA--TCGTAGGACTCGCTGCGAGCAAG	789		
Db	799	TCCTTTCACCTTAAGTGCAGCAGCTGCAAGCGCCCA--TCGTAGGACTCGCTGCGAGCAAG	740		
Qy	790	TATGTGTCCTTTGAAGACGACACTGGCAGACACAACTGCTTCCTGCGCCGCTGTCT	849		
Db	739	TATGTGTCCTTTGAAGACGACACTGGCAGACACAACTGCTTCCTGCGCCGCTGTCT	680		
Qy	850	ACCTCCCTGCTGGGCCAGGCTTCGTACCGGATGGAGACCAAGTCTGCCAGGGCTGT	909		
Db	679	ACCTCCCTGCTGGGCCAGGCTTCGTACCGGATGGAGACCAAGTCTGCCAGGGCTGT	620		
Qy	910	AGCCAGCAGGGCCCTTAAGCCAGGCTCTGGACCCAGGCTTTCCCATACACAGCGGCCA	969		
Db	619	AGCCAGCAGGGCCCTTAAGCCAGGCTCTGGACCCAGGCTTTCCCATACACAGCGGCCA	560		
Qy	970	GGACTGTGGCTCTCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAATG	1029		
Db	559	GGACTGTGGCTCTCTTTTCTAAACACCTCTGGGACTCAGCTCCCGCCGCAAAAATG	500		
Qy	1030	GGTCTCTCTGGGCTCCAGGATTTCTCCCACTCCAGATCCCAAACTGGTACTCC	1089		
Db	499	GGTCTCTCTGGGCTCCAGGATTTCTCCCACTCCAGATCCCAAACTGGTACTCC	440		
Qy	1090	TGACCCAGGCCCCCAATCTCTGGGCTTTACAGAGCTTCCATGATGATCAAGCCCTCCCA	1149		
Db	439	TGACCCAGGCCCCCAATCTCTGGGCTTTACAGAGCTTCCATGATGATCAAGCCCTCCCA	381		
Qy	1150	CACCTGACATCCAGAAATTCACCTCTCCCTGAGTCTGGTTCCTCCAGACTGAGTCTCT	1209		
Db	380	CACCTGACATCCAGAAATTCACCTCTCCCTGAGTCTGGTTCCTCCAGACTGAGTCTCT	321		
Qy	1210	CCCCAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCCC	1269		
Db	320	CCCCAAATCAGGGCTCTAGACCCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCCCCC	261		
Qy	1270	TAAATCTAGACTTCTCTTTATAGGTTTCAGGTTCTCTATGGGTGCTCGGGAAGTCTTG	1329		
Db	260	TAAATCTAGACTTCTCTTTATAGGTTTCAGGTTCTCTATGGGTGCTCGGGAAGTCTTG	201		
Qy	1330	AAAGTGACATGTTCTCAGGCTTGACCTGCCGCCACCCCATCCCGCGGTTGAGGCTGTGG	1389		
Db	200	AAAGTGACATGTTCTCAGGCTTGACCTGCCGCCACCCCATCCCGCGGTTGAGGCTGTGG	141		

QY	1390	GGCAGCAGATCAGGAGCCCACTGATTAAGGGGCCCTAGGGTACAGGTGCTGCCAGCAGG	1449
DB	140	GGCAGCAGATCAGGAGCCCACTGATTAAGGGGCCCTAGGGTACAGGTGCTGCCAGCAGG	81
QY	1450	TCGCCACCGAGTGCTCTTCATTTTATTTTCAGCTCCATTTTGGCCCATAGATGGCAGAGG	1509
DB	80	TKCKCWCAGAGTGCTCTTCATTTTATTTTCAGCTCCATTTTGGCCCATAGATGGTCAAGG	21
QY	1510	G 1510	
DB	20	G 20	
RESULT 13			
LOCUS	AL551558/c		
DEFINITION	AL551558 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA	1116 bp	mrna linear EST 31-MAY-2003
ACCESSION	AL551558		
VERSION	AL551558.2	GI:31273374	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
COMMENT	On Feb 15, 2001 this sequence version replaced gi:12889621.		
FEATURES	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by Life Technologies, a division of		
	Invitrogen. This sequence belongs to sequence cluster 888.f For		
	more information about this cluster, see		
	http://www.genoscope.cns.fr/		
	cgi-bin/cluster.cgi?seq=CS0DI063CB12NP1&cluster=888.f. Contact :		
	Feng Liang Email : fliang@lifetech.com URL :		
source	http://fulllength.invitrogen.com/Invitrogen Corporation 1600		
	Paraday Avenue Genoscope sequence ID : CS0DI063CB12NP1.		
	Location/Qualifiers		
	1..1116		
	/organism="Homo sapiens"		
	/mol_type="mrna"		
	/db_xref="taxon:9606"		
	/clones="CS0DI063YD23"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"		
BASE COUNT	primer. Five prime end enriched, double-strand cDNA was		
	digested with Not I and cloned into the Not I and EcoR V		
	sites of the pCMVSPORT 6 vector. Library was normalized."		
	235 a	259 c	337 g 226 t 59 others
ORIGIN			
Query Match	45.4%; Score 894.2; DB 9; Length 1116;		
Best Local Similarity	94.0%; Pred. No. 9.2e-217;		
Matches 913;	Conservative 26; Mismatches 27; Indels 5; Gaps 2;		
QY	548	AGAACAAAGTTGCTCTCGCTGCGCCGCTGCACAGACGCTCACACAGGTGGAGTGA	607
DB	984	AGAGAAAGATGTTGTCCKYCGCCCGCTTSMAGAAAGACGCTCACACAGGTGGAKTGA	925
QY	608	CATACCTGTATCAGCCCTGGCATCGAGAATCTCTGGTCTGTACCGGATGCCAGACGCC	667
DB	924	CATACCTGTATCAGCCCTGGCATCGAGAATCTCTGGTCTGTACCGGATGCCAGACGCC	865
QY	668	TGGCAGGCGCAGACTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTTTGGAG	727
DB	864	TGGCAGGCGCAGACTTCACCTCCCGGGATGAAGATCCCTACTGTGTGGCTGTTTGGAG	805

QY	728	AACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATCTGATAGACTCGGTGGAGGCA	787
Db	804	AACTCTTTGCACCTAAGTGCAGCAGCTGCAAGCGCCCATC---GACTCGGTGGAGGCA	749
QY	788	AGTATGTGTCCTTTGAAGACCCGACACTGGCACCACAACTGCTCTCTCGCGCCGCTGCT	847
Db	748	AGTATGTGTCCTTTGAAGACCCGACACTGGCACCACAACTGCTCTCTCGCGCCGCTGCT	689
QY	848	CTACTCTCCCTGTGTGGCCAGGCGCTTCGTACCGGATGGAGACCAAGTGTCTCTGCCAGGCT	907
Db	688	CTACTCTCCCTGTGTGGCCAGGCGCTTCGTACCGGATGGAGACCAAGTGTCTGCCAGGCT	629
QY	908	GTAGCCAGGACGGCCCTAAGCCAGGCTCCTTGACCCAGGCTTTCCCATACCAACGGGCC	967
Db	628	GTAGCCAGGACGGCCCTAAGCCAGGCTCCTTGACCCAGGCTTTCCCATACCAACGGGCC	569
QY	968	CAGGACTGTGGCTCTTTTCTAAACCACTCTCGGACTCAGCTCCCGCCCAAAAAA	1027
Db	568	CAGGACTGTGGCTCTTTTCTAAACCACTCTCGGACTCAGCTCCCGCCCAAAAAA	509
QY	1028	TGGGTCTCTCTTCTGGGCTCCAGGATGTCTCTCCCGACTCCAGCATCCCAAACTGGTACTC	1087
Db	508	TGGGTCTCTCTTCTGGGCTCCAGGATGTCTCTCCCGACTCCAGCATCCCAAACTGGTACTC	449
QY	1088	CCTGACCCAGGCCCCCAATCCTCTGGGCTCTTACAGAGCTCCATGATCAAGCCCTCC	1147
Db	448	CCTGACCCAGGCCCCCAATCCT--GGCTCTTTACAGAGCTCCATGATCAAGCCCTCC	390
QY	1148	CACACTGGACTCCAGAAATTCACCTCTCCCGTCTGAGTCTGGGTTCCGAGCTGAGTCTCT	1207
Db	389	MACACTGKACTCCAGAAATTCACCTCTCCCGTCTGAGTCTGGGTTCCGAGCTGAGTCTCT	330
QY	1208	CTCCCAAAATCAGGGCTCTAGACCGAGCCCTCCAAACCTGGACTCTGGGACTTAGGCC	1267
Db	329	CTCCCAAAATCAGGGCTCTAGACCGAGCCCAAAACCTGGACTCTGGGACTTAGGCC	270
QY	1268	CCTTAAATCTAGACTTCTCTTTATAGGTTTTCAGGCTCTCTATGGTGTCTGGGAAGTCT	1327
Db	269	CCTTAAATCTAGACTTCTCTTTATATTTTTCAGTCTCTCTATGGTGTCTGGAATCTCT	210
QY	1328	TGAAGTGGACTGTTCTCAGGCTTGACCTGCGCCCGACCCCATCCCGCGGTTGAGGCTGTG	1387
Db	209	TGAAGTGGACTGTTCTCAGGCTTGACCTKCCCGACCCCATCCCGCGGTTGAGGCTGTG	150
QY	1388	GGGGCAGCAGATCAGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTCTGCCAGCA	1447
Db	149	GGGGCAGCAGATCAGAGCCCACTGATAAGGGGCCCTAGGGTACAGGGTCTGCCAGCA	90
QY	1448	GGTCCGCCAGGAGTGTCTTCTCATTTTATTTTTCAGCTCCATTTTGCCCATAGATGGCAGA	1507
Db	89	GGTCCGCCAGGAGTGTCTTCTCATTTTATTTTTCAGCTCCATTTTGCCCATAGATGRACAGA	30
QY	1508	GGGGTGAGATT 1518	
Db	29	NANNTNNAATT 19	
RESULT 14			
BX445770/c			
LOCUS	BX445770	1201 bp	mRNA linear EST 15-MAY-2003
DEFINITION	clone CS0DH006Yall 3-PRIME, mRNA sequence.		
ACCESSION	BX445770		
VERSION	BX445770.1	GI:30774361	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1201)		
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		

JOURNAL COMMENT		Unpublished	
Contact: Genoscope		Genoscope - Centre National de Sequencage	
BP 191 91006 EVRY cedex - France		Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr	
Library was constructed by Life Technologies, a division of		Invitrogen. This sequence belongs to sequence cluster 888.f For	
more information about this cluster, see		http://www.genoscope.cns.fr/	
cqi-bin/cluster.cgi?seg=CS1DH002ZC03NP1&cluster=888.f. Contact :		Peng Liang Email : fliang@lifetech.com URL : Invitrogen Corporation 1600	
http://fulllength.invitrogen.com/		Faraday Avenue Genoscope sequence ID : CS1DH002ZC03NP1.	
Location/Qualifiers		1. .1201	
FEATURES		source	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="CS0DH006Yall"	
		/tissue_type="T CELLS (JURKAT CELL LINE)"	
		/cell_line="JURKAT CELL LINE"	
		/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"	
		/note="Vector: pCMVSPORT.6; 1st strand cDNA was primed	
		with a NotI-oligo(dT) primer. Five prime end enriched, into	
		double-strand cDNA was digested with Not I and cloned into	
		the Not I and EcoRV sites of the pCMVSPORT 6 vector.	
		Library was not normalized."	
BASE COUNT	259 a	293 c	349 g 229 t 71 others
ORIGIN			
Query Match	45.4%;	Score 894.2;	DB 13; Length 1201;
Best Local Similarity	90.6%;	Pred. No. 9.3e-217;	
Matches	969;	Conservative 32;	Mismatches 64; Indels 5; Gaps 5;
QY	461	GCTTCTGTGACGTGGCTGTGAACAGCCACTGGGCTCCCGTCTTTTGTGCGCGACAAG	520
Db	1068	KFTTCTTGGGGRADGKTTKAAAVCAAHGGKTBCCGKTGTGTTTSCCCAAAAGG	1009
QY	521	GTGCTCACTACTGCGTCCCTGTATGAGAACAAAGTTTGTCTGCTGCGCGCGCTGCA	580
Db	1008	GGGTTAAACACTGCKKCCCGGTTTARGAACARTTTKCCCCCTTCGSSCCGSGSCR	949
QY	581	GCAAGACGCTACACAGGCTGGAGTGACATCCGATGATCAGCGTGGCATCGAGATGTC	640
Db	948	GCAAAAGGTTAAACAGAGGGGAATTAACATACCGTGTTMAGCCGTGGCATCGARATTC	889
QY	641	TGCTGTACCGATGCCAGCCAGCCCTGGCAGGCGAGTTCACCTCCCGGATGAAG	700
Db	888	T-GTTTGTACCGGATGCCAACGCCCCCTGGCAGGCGAGTTCCTCCCGGGATGAAG	830
QY	701	ATCCCTACTGTGTGCGCTGTTTGGAGAACTCTTTGCACCTAAGTGACAGCTGCAAGC	760
Db	829	ATCCCAWACTTTTGGCTGTTTGGAGAACTCTTGCACCTAAGKGCDCAGCTGCAAGC	770
QY	761	GCCCCATCGTAGGACTCGGTGGAGGCAAGTATGTCTCTTTTGAAGACCGACACTGGC	820
Db	769	GCCCCATCGTAGAAGTCTGTGAGGCAAGTCTTGTCTCTTGAAGACCGACACTGGC	710
QY	821	ACAAGTCTTCTCTGCGCCGCTGCTCTACCTCCCTGCTGGCGCCAGGCTTCGTACCG	880
Db	709	ACAAGTCTTCTCTGCGCCGCTGCTCTACCTCCCTGCTGGCGCCAGGCTTCGTACCG	650
QY	881	ATGGAGACCAAGTGTCTGCCAGGCTGTAGCAGGCGCCCTAAGCAGGCGCTCTG	940
Db	649	ATGGAGACCAAGTGTCTGCCAGGCTGTAGCAGGCGCCCTAAGCAGGCGCTCTG	590
QY	941	GACCAAGGCTTTCATACACGCGCCAGGACTGTGGCTCTCTTTTAAACACACTCTG	1000
Db	589	GACCAAGGCTTTCATACACGCGCCAGGACTGTGGCTCTCTTTTAAACACACTCTG	530
QY	1001	GGACTCAGCTCCCCCGCC -AAAAAATGGCTCTCTCTGGCTCCAGGATTCCTCC	1059
Db	529	GAACTCAGCTCCCCCGCCAAAAAATGGGCTCTCTCTGGCTCCAGGATTCCTCC	470





Db           ||| | ||||| | | |||||  
          984. CCGGCCAAAAAATGGGCCTTCT 1009

Search completed: August 27, 2003, 14:11:44  
Job time : 2843 secs